**DOCUMENT RESUME**

ED 151 366 95 TM 006 411

AUTHOR Collet, LaVerne S.
TITLE FEHR-PRACTICUM: Information Bank for Project Head Start. Problems 4-8.
INSTITUTION Michigan Univ., Ann Arbor.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
PUB DATE [70]
NOTE 175p.; For related documents, see ED 059 630, 131 055, TM 006 408, 409, and 957; Problem 5 of the original document has been removed; Con Contact LeVerne S. Collet, 4218A School of Education Building, University of Michigan, Ann Arbor, Michigan 48104 for information; Some parts may be marginally legible due to print quality.

EDRS PRICE MF-$0.83 HC-$8.69 Plus Postage.
DESCRIPTORS *Annotated Bibliographies; Bus Transportation; College Mathematics; Compensatory Education Programs; Educational Programs; Elementary Secondary Education; Higher Education; *Integration Methods; Preschool Education; Program Evaluation; *Remedial Mathematics; *Research Reviews (Publications); Research Tools; Student Evaluation of Teacher Performance; Teacher Evaluation; Test Construction; *Test Reviews
IDENTIFIERS *Project Head Start

ABSTRACT

The Formative Evaluation and Heuristic Research (FEHR) PRACTICUM, a computerized game, gives practical experience in such program evaluation tasks as problem definition, operationalizing objectives, field study, budgeting, proposal writing, data analysis, and interpretation--without the expense and time commitments required by most research. The players, graduate students or practicing professionals, design and conduct an experiment comparing the effectiveness of available educational programs on students in a mythical school district. As a "research assistant," the computer can identify individuals or groups of students by category (test score, grade, sex), apply educational programs, and administer and score tests which determine the effects of a program over time. There is no predetermined decision; rather, the results of each team are critically evaluated and methodological implications are discussed. Available problems study from preschool to college level populations; this document describes empirical evaluations related to Head Start programs, remedial mathematics for college students, busing to achieve racial integration, and validation of a teacher questionnaire. Each problem has its own Information Bank, which is a simulated library that helps players to define their experiment by examining real-life research. Each bank is divided into four sections: (1) abstracts of articles directly related to the problem, (2) abstracts of auxiliary references, (3) descriptive and normative data on each of the standardized tests available in the problem, and (4) test critiques. (CP)

ED151366

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

LeVerne S. Collet

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

FEHR-PRACTICUM
INFORMATION BANK
PROBLEMS 4-8
COPY 2

TM006 411

INFORMATION BANK FOR

PROJECT HEADSTART

These materials are meant to be used as abstracts or summaries, highlighting the major points of the articles or presenting only the most relevant normative information. They should be referred to in the same manner as one would refer to a journal such as Psychological Abstracts. More complete information on these sources, or sources not included here, may be found in your library.

This information bank is divided into four sections, as follows: (1) References directly related to this problem; (2) Auxiliary References; (3) Descriptive and normative data about variables; and (4) Test critiques, validation studies, etc.

Academic Preschool: Champaign, Illinois; U.S. Department of Health, Education and Welfare, OE-37041.

The design of the academic preschool program was based on the rationale that children who are one or more years retarded in language, reading, or mathematics must learn at a rate which exceeds the learning rates for normal children in order to catch up. Fifteen behavioral objectives defined the specific performance criteria for the children to meet. The curriculum focused directly on these objectives ensuring that the child received an amount of exposure, practice, and correction sufficient to teach what was intended. Direct instruction, similar to that used in regular school, was employed as an alternative to the informal style of the traditional nursery.

The experimental children were 4 and 5 year olds, predominantly Negro, and of low socio-economic status. Their parents were usually unskilled or semi-skilled laborers; at least 30 to 40 percent were receiving some welfare assistance.

The program began in the fall of 1964 and data were collected through the spring of 1968. Each of three groups of approximately 15-20 children received the treatment for 2 years prior to their entrance into the first grade. The first group, Study I, participated in the preschool during academic years 1964-66; they completed the second grade in June 1968. The two subsequent groups, Studies II and III, were in the program from 1965-67 and 1966-68, respectively. The most recent data were collected from a follow-up study of groups I and II in the early elementary grades. Only Study II will be described here, since it was the only study for which there was both an experimental and comparison group.

The effectiveness of the program was indicated by the significant superiority of the experimentals over the controls in Stanford-Binet IQ gains over the 2-year period of instruction. In addition, upon completion of the 2-year preschool program the experimentals tested considerably above the first grade level in mathematics and language as measured by the Wide Range Achievement Tests.

Ball, S. and Bogatz, G.A. "The First Year of Sesame Street: An Evaluation." Princeton, New Jersey: E.T.S., 1970.

In its first season of 26 weeks, Sesame Street showed that television can be an effective medium for teaching 3-5-year-old children important, simple facts and skills, such as recognizing and labelling letters and numerals, and more complex higher cognitive skills, such as classifying and sorting by a variety of criteria. The ETS research results reveal that Sesame Street benefits children from disadvantaged inner-city communities, middle class suburbs, and isolated rural areas--all the groups studied in this evaluation.

First, children who watched the most learned the most. The amount of learning that took place--that is, the gains a child showed between being tested for certain skills before watching Sesame Street and being tested for the same skills after--increased in relation to the amount of time the child watched the program.

Second, the skills that received the most time and attention on the program itself were, with rare exceptions, the skills that were best learned. An analysis of the content of the show revealed, for example, that more time (13.9 percent) was devoted to letter-related skills than to any other single subject; it was in the areas of letters and numbers that the children's gains were the most dramatic. In addition to acquiring skills that were directly and deliberately taught, it appears that there was some transfer of learning, that some children learned to do things--such as recognize full words or write their own names--which were not taught on the program.

The 3-year-old children gained the most; 5-year-olds gained the least. That is, 3-year-old children who viewed the show a great deal had higher attainments at post-test than those 4- and 5-year-olds who viewed the show less, even though the younger children scored lower at pretest than the older children. This finding has important implications for education in general, for it suggests that 3-year-old children are able to learn many skills that have traditionally been introduced at later ages.

A similar phenomenon appeared with advantaged and disadvantaged children. Although the disadvantaged children started out with considerably lower achievement scores on the skills being taught, those who watched a great deal surpassed the middle class children who watched only a little. It thus appears that such television programs can reduce the distinct educational gap that usually separates advantaged and disadvantaged children even by the time they enter first grade.

Bereiter, Carl and Engelmann, Siegfried. "Language Learning Activities for the Disadvantaged Child." Anti-Defamation League of B'nai B'rith, New York, 1966.

"It appears that in learning any of the modern languages from educated, articulate parents, a child learns certain rules about how language operates and about what can be done with it that are readily transferred to any other language that he learns. But a child who grows up in a social group that for generations has known only poverty and unskilled employment, where formal education is little known, and where the teaching that is done is done by outsiders, does not learn these language rules, even if the language he learns is fundamentally the same as the language of those who will teach him. He may have learned language rules that are adequate for expressing his wants, for following concrete instructions, for expressing feelings, and possibly for telling stories; but he has not learned the language rules that are necessary for defining concepts, for drawing inferences, for asking questions, and for giving explanations. He has not learned enough about the detailed character of words and their sounds to be able to understand his own language when it is written down a word at a time...By his inability to make full use of language as a tool in learning and thinking, he is prevented from taking full advantage of the opportunities for education and advancement that are at last being made available to him."

Bernstein, Basil. "Language and Social Class." British Journal of Sociology, 1960, 11, pp. 271-276.

Matched groups of 61 working class and 45 middle-class high-school age English children were tested on the Raven's Progressive Matrices Test (a non-verbal intelligence test) and the Mill-Hill Vocabulary Scale. The middle class group achieved scores in the average to superior range, their scores on the two tests being very similar. The working class group, however, showed large discrepancies between the two scores: although none of the working class children scored over 109 on the vocabulary test, 36 of them scored over 109 on the Matrices Test and 11 scored 120 or above. (Thirty-four of the middle class children scored over 109 on the Matrices Test, and 24 scored 120 or above.)

Bloomington, Indiana, Diagnostically Based Curriculum, U.S. Department of Health, Education and Welfare, OE-37024, 1969.

The purpose of this study was to develop and evaluate a diagnostically based curriculum for disadvantaged preschool children. The diagnosis attempted to identify the child's learning deficits in language, concept, and fine motor development so that specific remediation procedures could be applied individually. With the exception of specifically developed diagnostic language and fine motor lessons, the experimental curriculum included many of the kinds of activities found in regular preschool and kindergarten programs. However, activities such as music, art, free play, and story-telling were not used as ends in themselves, but as vehicles for the development and remediation of specific cognitive, psycho-motor, and social behaviors.

The student population consisted of 139 Appalachian white and four Negro children from families of the lowest socio-economic class as determined by the Warner-Meeker-Eells Index of Status Characteristics. From those children who met the socio-economic criteria, only 5-year old children who scored between 50-85 on the 1960 Stanford-Binet L-M Intelligence Scale were selected.

Three studies with similar design were completed in 3 consecutive years (1964-67) preceded by a 1-semester pilot study. A longitudinal study of the children as they proceeded through various public schools was also conducted. In each of the three studies approximately 15 children were placed in either an Experimental Preschool (EPS), Kindergarten Contrast (KC), or At Home Contrast (AHC) group. The EPS children received the diagnostic treatment; the KC children received a traditional kindergarten treatment; the AHC received no treatment. Although the original intent was to replicate the basic study twice to accumulate larger numbers of subjects within each treatment, the three studies were not precise replications. The diagnostically developed curriculum was modified each successive year by incorporating experiences gained from the previous year.

The diagnoses were made on the basis of Stanford-Binet scores and direct observations of teachers and project staff. Measurement of gains in language development, IQ, and motor development were made with five standardized tests. Results for intelligence and language development favored the EPS group. Results in Fine and Total Motor skill equally favored the EPS and KC groups over continued residence in the home.

Hess, Robert D. and Shipman, Virginia. "Early Blocks to Children's Learning." Children, 1965, 12, No. 5.

The authors investigated "linguistic and interactional styles" of 160 Negro mothers and their 4-year-old children, selected from four socio-economic levels. Each mother was taught three simple tasks and then told to teach them to her child. Great social class differences appeared in mothers' teaching skills, as well as in children's performances. The authors report that the lower-class mothers failed to provide their children with the essential information needed to solve problems, and often did not even explain what the child was supposed to do. It is concluded that "the meaning of deprivation is a deprivation of meaning...(enrichment) programs must not merely teach the child new words, but must show the child how ideas and events are related to one another."

Infant Education Research Project. U.S. Department of Health, Education and Welfare, OE-37033, Washington, D.C.

"Tutors provided intellectual and verbal stimulation to children, from the time the children were age 15 months until they were age 36 months. Instruction was on a one tutor, one child basis, 1 hour per day, 5 days per week, in the child's home.

The children were Negro males. They were drawn from homes that met two of the following three criteria: (1) the mother's formal education was less than 12 years; (2) the mother's work history was of unskilled or semi-skilled employment; and (3) the family income was $5000 or less. Fathers were absent in many of the homes; the families were of lower socio-economic status. All lived in deteriorated areas of Washington, D.C. An attempt was made to choose the participants from relatively stable homes, not so noisy or overcrowded as to interfere with the home tutoring sessions which were central to the project. Children were not selected if they displayed serious sensory or neurological problems."

"Training of tutors and selection of children began in September 1965 and tutoring began 2 to 3 months later. There were 28 children in the experimental group and 30 in a control group."

Various tests of intelligence were used as a measure of the effectiveness of the program. The Bayley Infant Scales of Development were used as a pretest of all children at 14 months. Later the Stanford Binet, and the Peabody Picture Vocabulary Tests were used. There were no significant differences between the control and experimental groups at 14 months. At 36 months the experimental group was significantly better on the Stanford Binet and the PPVT control group.

Kagan, Jerome and Moss, H.A. "Parental Correlates of Child's IQ and Height: A Cross-Validation of the Berkeley Growth Study Results," Child Development, 1959, 30, 325-332.

"It is suggested that maternal education is a good predictor of IQ score in both sexes because level of education may be a sensitive index of the mother's concern with her child's intellectual development. That is, level of maternal education may be positively correlated with (a) the degree to which the mother rewards and encourages intellectual development in the child and (b) the degree to which the child perceives that the mother values intellectual skills and activity. Since the mother's contact with the child during the preschool and early school years tends to be more extensive than that of the father, it is plausible to expect that maternal concern with intellectual development will have a stronger influence on the child's desire to master intellectual skills than similar concern on the part of the father. The fact that maternal education predicts girls' IQ better than boys' IQ might be anticipated since a daughter is more apt than a son to identify with the values of the mother. Previous research by the authors tends to support these hypotheses for ratings of maternal concern with intellectual development during the child's first three years showed a significant, positive relationship, for girls only, with both achievement fantasy and increase in IQ during the years 6 through 10."

Kraft, I.; Fuschillo, J.; and Herzog, E. "Prelude to School: An Evaluation of an Inner-City Preschool Program." U.S. Department of Health, Education and Welfare, 1968.

The purpose of the Project was to determine whether a 2-year experience in a traditional middle-class nursery school, plus parent education activities, could help offset the difficulties which low-income, inner-city children experience when they enter elementary school.

The families from which both the nursery and comparison groups were drawn were similar in salient socio-economic variables, such as family yearly income (median annual income was approximately $3,500 for both groups), income per child, extent of welfare assistance, number of children in home, and mother's education.

To evaluate the impact of the nursery school program, both nursery and comparison group children were tested before the former group's entrance into preschool, at the end of the first year of preschool and at the end of the second year.

The Stanford-Binet was administered on all three test rounds. The Peabody Picture Vocabulary Test, portions of the Merrill-Palmer, and two subtests from the Illinois Test of Psycholinguistics were given in two of the three test rounds, as a supplement to the Stanford-Binet.

The average Stanford-Binet IQ of the nursery-group children prior to their preschool experiences was 82.7, with a range of 53-113; the comparison group's average was 84.6, with a range of 65-119. In both the nursery and comparison groups, mean IQ scores showed a consistent, though not statistically significant relation to family SES level, with the children from the highest level, within this low-income group, receiving the highest scores, and so on down the line.

After 2 years of preschool, the nursery group's average IQ was 97.4, a gain of 14.7 points. The nursery-group boys, whose initial average IQ was only 77.3 (which may have been the result of a sampling bias) gained 19.6 points, while the girls whose initial IQ averaged 86.3, gained 11.4 points. The comparison group in the same 2-year period made an average gain of 4.1 points.

Raph, J.; Lieberman, D.; and Pascale, P. Influence of a Piaget-Oriented Curriculum on Intellectual Functioning of Lower-Class Kindergarten Children. American Educational Research Association, New York, 1971.

"Intellectual functioning of a lower-class kindergarten class exposed to the introduction of a Piaget-oriented curriculum was compared with intellectual functioning of one comparison lower-class and one comparison middle-class group of children taught in a traditional, activity-centered kindergarten program. Change scores between the beginning and end of a school year were significantly higher for the experimental lower-class group on one standardized measure and on several Piagetian tasks. Main features of the Piaget-oriented curriculum and implications of the results are discussed."

"What is interpreted as the strongest evidence that the innovative features of the kindergarten were, indeed, effective, is shown most clearly on the three most difficult tasks in the Piagetian battery, Reverse Linear Ordering, Seriation II, and Conservation. In these instances, the experimental group made significant gains, the comparison lower- and middle-class groups did not. This outcome may be said to reflect something of Piaget's position that pedagogical interventions can accelerate or complete spontaneous development, but cannot change the order of constructions (Piaget, 1970)."

Schaefer, E.S. Intellectual Stimulation of Culturally Deprived Infants. Laboratory of Psychology, National Institute of Mental Health, 1968.

Method: An experimental group of 31 Negro male infants and a control group of 33 Negro male infants were selected from door-to-door surveys of two lower socio-economic neighborhoods in Washington, D.C. Although experimental and control cases were chosen from different neighborhoods to avoid contamination, children from the two neighborhoods have comparable readiness scores at school entrance. Only families that had agreed to participate in either the experimental or control group were included. Families that did not meet two of the three criteria--family income under $5000, mother's education under 12 years, and mother's occupation, if any, either unskilled or semi-skilled--were excluded as were homes judged to be too crowded or unsuitable for home tutoring. Comparisons between the groups revealed only small differences, many of which favored the control group, on family variables that might be expected to influence the child's intellectual development.

Tutors visited the home of each experimental infant for an hour a day, five days a week beginning when the infant reached 15 months of age and continuing through 36 months. The educational program included major emphases on the development of positive relationships, the provision of varied experiences, and verbal stimulation of the infants. Participation of the mother and of other family members in the education of the infant was encouraged but not required.

Comparable groups of experimental and control male infants were selected and administered an infant mental test at 14 months of age. Home tutoring was provided for the experimental group beginning at 15 months. Both groups were retested at 21, 27, and 36 months. Major findings of the project are that both supplementary tutoring and the adequacy of maternal care are significantly related to child behavior and mental test scores at 36 months of age.

Sontag, M.; Sella, A.P.; and Thorndike, R.L. "The Effect of Headstart Training on the Cognitive Growth of Disadvantaged Children." Journal of Educational Research, 62(9), 1969, 387-389.

To measure the effects of Headstart on the cognitive development of preschool children, forty-three pairs of Headstart children were compared with children of similar ages who had been registered for Headstart but had not as yet begun the program. The assumption made in the selection of the latter group for comparison purposes was that the family background and other variables that make for a particular child being enrolled in this voluntary program would be controlled. Subjects were given the Stanford-Binet Intelligence Test and the Caldwell-Soule Preschool Inventory. No significant differences ($p = > .05$) in IQ were found. Significant differences (approximately one-half of a standard deviation) were found on the Caldwell-Soule Preschool Inventory.

Weikert, D. The Perry Preschool Project in Ypsilanti, Michigan, 1969, OE-37035.

The Perry Preschool Project was aimed at assessing the longitudinal effects of a two-year program consisting of a daily 3-hour cognitively oriented nursery, a weekly 90-minute home visit, and less frequent group meetings of the pupils' parents.

The students were 3- and 4-year old Negro disadvantaged and functionally retarded children, whose pretest scores on the Stanford-Binet Intelligence Scale were not above 85.0.

The program described here operated from September 1962 until June 1966. Approximately 24 children participated in the preschool annually. Half of these were 3 years of age, the other half were 4-year olds. Each of these age groups was designated as a "Wave"; consequently, two different "Waves" participated each year.

For each group or "Wave" of children entering the experiment, the Stanford-Binet Intelligence Scale, the Leiter International Performance Scale, and the Peabody Picture Vocabulary Test were used at the start of the experiment to test the intellectual ability of the children. Few significant differences between the experimentals and the controls were noted at this time. These and other tests were also used to measure intellectual ability later in the program. The benefits claimed for the program are those measured by the California Achievement Test in reading, language, and mathematics. These tests were given at the end of first grade and again at the end of second grade. The results showed significant gains for the experimentals over the controls.

Wellman, Beth L. "I.Q. Changes of Preschool and Non-preschool Groups During the Preschool Years: A Summary of the Literature." Journal of Psychology, 1945, 20, 347-368.

"The IQ changes of preschool (nursery school) and non-preschool groups during the preschool years, as found in the literature, have been summarized in tabular form arranged according to the scale utilized. Approximately 50 articles giving such changes were located. There were 22 preschool groups and 14 non-preschool groups tested on revisions of the Binet scale, showing changes ranging from slight negative to large positive amounts. Fifty percent of the preschool groups gained six or more points, while only 14 percent of the non-preschool groups made similar gains; 71 percent of the non-preschool group changes were lower than a two-point gain. On the Merrill-Palmer scale, seven preschool groups and four non-preschool groups were tested. All but one preschool group and only one of the non-preschool groups gained six or more points...For the preschool children tested on the Binet Scales the mean of group changes was 5.4 points and the mean change per child for the total of 1,537 children was also 5.4 points; for non-preschool children tested on these scales the mean of group changes was 1.1 points and the mean change per child for the total changes was 1.1 points and the mean change per child for the total of 597 children was 0.5 point. For the preschool children tested on the Merrill-Palmer scale the mean of group changes was 13.0 points and the mean change per child for the total of 267 children was 14.5 points; for the non-preschool children the mean of group changes was 6.2 points and the mean change per child for the total of 73 children was 6.7 points.

deHirsch, Katrina; Jansky, Jeannette Jefferson; and Langford, William S. Predicting Reading Failure, A Preliminary Study of Reading, Writing, and Spelling Disabilities in Pre-School Children. Harper & Row Publishers, New York, 1966, 144 pp.

Twenty years of clinical experience with intelligent, but educationally disabled children, whose learning drive has become severely damaged, has convinced us that many of these children would not have required help had their difficulties been recognized at early ages. Early identification would have obviated the need for later remedial measures.

This study has attempted to develop techniques for the early identification of children who, at kindergarten age, seemed to present a specific pattern of dysfunctions, reflecting an underlying developmental lag.

A battery of ten tests was finally chosen as the best identifier of children who would fail in reading. This included: Pencil Use, Bender Gestalt Test, Wepman Auditory Discrimination Test, Number of Words Used in a Story, Categories, Horst Reversals Test, Gates Word Matching Test, Word Recognition I and II, and Word Reproduction.

The children whose reading achievement at the end of second grade comprised the bottom 15% of the sample were termed "Failing Readers." As a group, they were markedly hyperactive, distractible, impulsive and disinhibited. Their auditory-perceptual and oral language tools were inferior to those of the other subjects. The "slow starters," those who showed great improvement from the end of 1st to 2nd grade, didn't show these traits. The "superior achievers," the top 15% of the sample, were described as physically and emotionally mature.

Children's developmental rhythm varies widely. Recognition of and respect for these variations are crucial at a time when society places increased pressures for early achievement on both children and parents. Such recognition implies the taking of active educational measures geared to the child's individual needs at his particular developmental level.

DiCarlo, L.M. "Critique of Audiometers." In O.K. Buros (ed.) The Sixth Mental Measurement Yearbook. Highland Park, New Jersey; The Gryphon Press, 1965, 940.

"Beltone Model 9-C. The Beltone 9-C is a basic audiometer for screening purposes. It is portable, durable and easy to manipulate. The instrument has proved to be stable and drift free. Purity of tone and accuracy of output levels meet the criteria for screening audiometers. It has a slide wire attenuator and is easy to calibrate. The output limit is 100 db."

ELEMENTARY SCHOOL INFORMATION - FAIR CITY

| INDEX | NAME | ADDRESS & TELEPHONE | OPENED | ADDITION | INFORMATION FOR 9-72 Teaching Stations | Classrooms in Use | Portables | Enrollment | Teachers | Site (Acres) | PER PUPIL COST |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | L.J. Horner | 2670 Tyler St. 936-3566 | 1957 | | 12 | 11 | 0 | 367 | 12 | 10 | $605 |
| 2 | P.P. Pumpkineater | 2560 Johnson Ave. 158-8565 | 1941 | 1960 | 27 | 26 | 2 | 808 | 27 | 11 | $612 |
| 3 | Knaveofhearts | 1608 Douglas St. 693-0266 | 1953 | 1965 | 21 | 20 | 0 | 630 | 21 | 9 | $625 |
| 4 | Jackandjill | 600 Hamilton St. 293-0622 | 1923 | | 25 | 24 | 1 | 786 | 25 | 8 | $597 |
| 5 | H.D. Dock | 2775 Burr Drive 168-2791 | 1961 | | 21 | 20 | 0 | 651 | 21 | 9 | $618 |
| 6 | M. Goose | 1414 Jay Avenue 608-4686 | | | 22 | 20 | 0 | 660 | 21 | 10 | $621 |
| 7 | Lamb | 4250 Lee Street 904-0471 | 1959 | | 25 | 23 | 1 | 714 | 24 | 11 | $630 |
| 8 | L.M. Muffet | 2935 Grant Avenue 691-4700 | 1967 | | 16 | 14 | 0 | 455 | 15 | 11 | $627 |
| 9 | J. Sprat | 2135 Allen Lane 893-0265 | 1940 | | 21 | 17 | 0 | 457 | 20 | 15 | $682 |
| 10 | Calicocat | 3540 Darwin Ct. 719-2663 | 1955 | | 28 | 25 | 1 | 772 | 27 | 17 | $701 |
| 11 | T. Blindmice | 800 Lafayette Dr. 888-6866 | 1970 | | 25 | 22 | 0 | 666 | 24 | 18 | $710 |
| 12 | H.D. Diddle | 825 Patton Ave. 821-6597 | 1922 | 1951 | 22 | 20 | 0 | 621 | 22 | 14 | $699 |
| 13 | O.K. Cole | 3800 Stowe St. 895-7771 | 1960 | | 26 | 24 | 1 | 762 | 25 | 12 | $660 |



(Continued)

| | | | | | INFORMATION FOR 9-72 | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| INDEX | NAME | ADDRESS & TELEPHONE | OPENED | ADDITION | Teaching Stations | Classrooms in Use | Portables | Enrollment | Teachers | Site (Acres) | PER PUPIL COST |
| 14 | Shaftoe | 344 Anthony Ave. 650-4665 | 1969 | | 27 | 24 | 1 | 766 | 26 | 15 | $671 |
| 15 | M.M. Contrary | 2250 Keyes Ct. 564-8760 | 1958 | | 19 | 15 | 0 | 339 | 18 | 18 | $705 |
| 16 | Inashoe | 920 Frost St. 561-6623 | 1965 | | 22 | 20 | 0 | 662 | 21 | 13 | $640 |
| 17 | S. Simon | 4377 Blackwell Ave. 112-3434 | 1956 | 1963 | 18 | 17 | 0 | 616 | 18 | 10 | $599 |
| 18 | Hubbard | 3550 King Drive 179-4150 | 1944 | 1952 | 16 | 15 | 0 | 554 | 16 | 11 | $590 |
| 19 | Sixpence | 2775 Wilson St. 178-0200 | 1952 | | 11 | 11 | 0 | 420 | 11 | 9 | $593 |
| 20 | B.B. Sheep | 2100 Harding St. 194-0620 | 1925 | | 13 | 12 | 0 | 441 | 12 | 7 | $580 |
| 21 | L.B. Blue | 2543 Lowell Rd. 197-3774 | 1962 | | 8 | 7 | 0 | 241 | 7 | 8 | $586 |

Robinson, H.M. "Visual Screening Tests for Schools." Elementary School Journal, 1953, 54, 217-222.

Visual Tests have been deemed important for schools to use for some time since good vision is certainly an aid if not a prerequisite for learning. The visual screening tests which ahve been developed are not to be used for diagnosis but only to identify children with a problem.

Researchers have done little to examine carefully the reliability or validity of such instruments or tests. This paper analyzes four such commercial batteries: The Keystone Visual Survey Test, the Massachusetts Vision Test, the Eames Eye Test, and the Ortho-Rater Test of Visual Efficiency. The subtests of each battery differ somewhat. The Ortho-Rater and the Keystone have the most subtests and test the most variety of visual skills. These two are also the most expensive. The batteries were given to 79 pupils in grades I-VIII and their validity was checked by comparing the test results with an opthalmological examination. The highest agreement between test result and opthalmological result was with the Ortho-Rater and the Massachusetts batteries. Considerably lower but still well correlated was the Keystone. The Eames battery did not correlate well at all.

The Reliability of these tests is not reported except for the Eames which is .93. A pilot reliability study was done with the Ortho-Rater and the reliabilities ranged from .68 to .82 with a mean of .77 on the test retest reliability.

The data here is preliminary but it points to the caution which should be used in selecting visual screening tests.

California Test of Mental Maturity (Short Form) (1 Form)

Purpose: To provide information about the functional capacities that are basic to learning, problem-solving, and responding to new situations.

Type of Test: Group intelligence

Range: Grades K-16

Time for Administration: 34-48 minutes

Cost: $4.20 per 35 tests, $.10 per scoring key

Subtests:
- Opposites
- Similarities
- Analogies
- Delayed Recall
- Numerical Values
- Number Problems
- Verbal Comprehension

Levels:
- 0 - Preprimary
- 1 - Primary
- 1H - Spans range between grades 3-4
- 2 - Elementary
- 2H - Spans range between grades 6-7
- 3 - Junior High
- 4 - Secondary
- 5 - Advanced

Scores: Scaled; Language IQ, Non-Language IQ, Total IQ

Nature of Test:

The California Short-Form of Mental Maturity consists of seven administrative test units, measuring aspects of mental ability, which contribute to a pattern of summary and derived scores (mental ages, intelligence quotients, standard scores, stanines, and percentile ranks) that are interpreted within a framework of inter and intra-individual differences. In addition to assessing the development of an individual or group with reference to national performance standards at each age level, the test results provide data as to the nature and potential of the abilities possessed by the individual. The Short-Form thus serves both survey and analytical purposes for educators, counselors, psychologists, and employers in a wide variety of testing situations.

California Short Form Test of Mental Maturity - Mean Intelligence Quotients & Standard Deviations of Mental Maturity Series, 1963 Revision & the Stanford-Binet Intelligence Scale Form LM 1960 Revision, Based on the Performance of the Sample Groups

| Short-Form Level | Type of CTMM IQ | IQ Means | | IQ SD's | |
|---|---|---|---|---|---|
| | | Cal. | Binet | Cal. | Binet |
| Level 0 | Total | 109.4 | 108.2 | 14.9 | 14.8 |
| | Language | 109.4 | | 16.6 | |
| | Non-Language | 107.6 | | 14. | |
| Level 1 | | Cal. | Binet | Cal. | Binet |
| | Total | 110.2 | 109.2 | 15.1 | 15.0 |
| | Language | 109.3 | | 15.5 | |
| | Non-Language | 109.8 | | 15.5 | |

| Short-Form Level | Type of CTMM IQ | IQ Means | | IQ S.D.'s | |
|---|---|---|---|---|---|
| Level 2 | | Cal. | Binet | Cal. | Binet |
| | Total | 103.7 | 106.7 | 14.2 | 13.8 |
| | Language | 104.5 | | 15.0 | |
| | Non-Language | 102.3 | | 14.2 | |
| Level 3 | | Cal. | Binet | Cal. | Binet |
| | Total | 105.5 | 106.5 | 15.0 | 14.9 |
| | Language | 104.9 | | 14.8 | |
| | Non-Language | 105.0 | | 15.5 | |
| Level 4 | | Cal. | Binet | Cal. | Binet |
| | Total | 108.9 | 111.8 | 13.9 | 14.3 |
| | Language | 106.1 | | 14.5 | |
| | Non-Language | 109.2 | | 14.0 | |
| Level 5 | | Cal. | Binet | Cal. | Binet |
| | Total | 113.9 | 116.4 | 14.5 | 14.9 |
| | Language | 112.1 | | 14.5 | |
| | Non-Language | 111.9 | | 14.1 | |
| All Levels | | Cal. | Binet | Cal. | Binet |
| | Total | 108.6 | 109.7 | 15.0 | 15.0 |
| | Language | 107.8 | | 15.4 | |
| | Non-Language | 107.6 | | 15.0 | |

California Test of Mental Maturity (Short Form)

| Variables | Reliability Short Form | Validity Coefficients of Correlation |
|---|---|---|
| Level 0 | | |
| I.Q. Total | .78 | .68 |
| Language | .70 | .61 |
| Non-Language | .59 | .56 |
| M.A. Total | .78 | .80 |
| Language | .70 | .74 |
| Non-Language | .59 | .71 |
| Level 1 | | |
| I.Q. Total | .83 | .70 |
| Language | .73 | .64 |
| Non-Language | .72 | .57 |
| M.A. Total | .83 | .89 |
| Language | .73 | .87 |
| Non-Language | .72 | .83 |
| Level 2 | | |
| I.Q. Total | .95 | .74 |
| Language | .94 | .70 |
| Non-Language | .89 | .61 |
| M.A. Total | .95 | .80 |
| Language | .94 | .76 |
| Non-Language | .89 | .69 |

| Variables | Reliability | Validity |
|---|---|---|
| Level 0 | Short Form | Coefficients of Correlation |
| Level 3 | | |
| I.Q. Total | .91 | .78 |
| Language | .90 | .77 |
| Non-Language | .81 | .65 |
| M.A. Total | .91 | .78 |
| Language | .90 | .77 |
| Non-Language | .81 | .63 |
| Level 4 | | |
| I.Q. Total | .92 | .71 |
| Language | .91 | .68 |
| Non-Language | .83 | .57 |
| M.A. Total | .92 | .76 |
| Language | .91 | .73 |
| Non-Language | .83 | .63 |
| Level 5 | | |
| I.Q. Total | .91 | .75 |
| Language | .90 | .75 |
| Non-Language | .82 | .56 |
| M.A. Total | .91 | .75 |
| Language | .90 | .75 |
| Non-Language | .82 | .57 |
| Combined | | |
| I.Q. Total | .91 | .73 |
| Language | .90 | .69 |
| Non-Language | .82 | .60 |
| M.A. Total | .91 | .96 |
| Language | .90 | .96 |
| Non-Language | .82 | .94 |

California Test of Mental Maturity (Critique)

See Also:

Anderson, H.E. and Leton, D.A. "Factor Analysis of the California Test of Mental Maturity." Educational and Psychological Measurements, 1964, 24(3), 513-523.

Norman, R.D. and Daley, M.F. "The Comparative Personality Adjustment of Superior and Inferior Readers." Journal of Educational Psychology, 1959, 50(1) 31-36.

California Test of Personality (2 Forms)

Purpose: To identify and reveal the status of certain highly important factors in personality and social adjustment usually designated as 'intangibles.'

Type of Test: Group personality

Range: Grades 1-adult (may also be used for kindergartin )

Time for Administration: 45-60 minutes

Cost: $3.50 per 35 tests

Subtests:
- Self-Reliance
- Sense of Personal Worth
- Sense of Personal Freedom
- Feeling of Belonging
- Withdrawing Tendencies
- Nervous Symptoms
- Social Standards
- Social Skills
- Anti-Social Tendencies
- Family Relations
- (School/Occupational) Relations
- Community Relations

Scores: Scaled for each subtest:

Personal Adjustment (Sum of first 5 subtests)

Social Adjustment (Sum of last 5 subtests)

Total Adjustment (Sum of personal & social adjustment)

Nature of Test:

The California Test of Personality has been designed to identify and reveal the status of certain highly important factors in personality and social adjustment usually designated as intangibles. These are the factors that defy appraisal or diagnosis by means of ordinary ability and achievement tests. Measurements of capacity, skill, and achievement, important as they are, do not constitute a complete picture of a functioning personality.

California Test of Personality - Means & Standard Deviations of T Scores on California Test of Personality Variables for Both Groups

| Variables | Inferior Readers | | Superior Readers | |
|---|---|---|---|---|
| | Mean | SD | Mean | SD |
| Personal Adjustment: | | | | |
| Self-Reliance | 46.90 | 6.77 | 52.28 | 7.27 |
| Sense of Personal Worth | 50.49 | 8.30 | 57.33 | 7.36 |
| Sense of Personal Freedom | 46.17 | 7.57 | 53.17 | 6.88 |
| Feeling of Belonging | 47.41 | 8.81 | 56.50 | 6.79 |
| Withdrawing Tendencies (Freedom from) | 51.32 | 7.84 | 59.38 | 6.96 |
| Nervous Symptoms (Freedom from) | 49.44 | 11.43 | 55.17 | 7.61 |
| Social Adjustment: | | | | |
| Social Standards | 43.95 | 8.98 | 53.83 | 7.36 |
| Social Skills | 49.17 | 9.84 | 54.36 | 10.19 |
| Anti-Social Tendencies (Freedom from) | 44.80 | 10.21 | 52.14 | 9.44 |
| Family Relations | 45.12 | 8.47 | 52.52 | 6.83 |
| School Relations | 50.05 | 9.79 | 57.71 | 7.14 |
| Community Relations | 43.07 | 8.72 | 50.43 | 5.22 |
| Total Adjustment: | | | | |

## California Test of Personality

| Reliability: Components: | Prim. r | Elem. r | Inter. r |
|---|---|---|---|
| Personal Adjustment | .83 | .93 | .93 |
| A. Self-Reliance | .73 | .64 | .70 |
| B. Sense of Personal Worth | .82 | .79 | .75 |
| C. Sense of Personal Freedom | .73 | .79 | .92 |
| D. Feeling of Belonging | .70 | .77 | .97 |
| E. Withdrawing Tendencies (Fdm.) | .71 | .83 | .83 |
| F. Nervous Symptoms (Fdm.) | .87 | .82 | .82 |
| Social Adjustment | .80 | .92 | .94 |
| A. Social Standards | .51 | .59 | .94 |
| B. Social Skills | .70 | .73 | .75 |
| C. Anti-Social Tendencies (Fdm.) | .82 | .77 | .86 |
| D. Family Relations | .82 | .77 | .92 |
| E. School Relations | .70 | .78 | .86 |
| F. Community Relations | .78 | .79 | .87 |
| Total Adjustment | .88 | .94 | .96 |

| Reliability Components: | Sec. r | Adult r |
|---|---|---|
| Personal Adjustment | .90 | .93 |
| A. Self-Reliance | .70 | .81 |
| B. Sense of Personal Worth | .77 | .66 |
| C. Sense of Personal Freedom | .84 | .87 |
| D. Feeling of Belonging | .91 | .87 |
| E. Withdrawing Tendencies (Fdm.) | .86 | .86 |
| F. Nervous Symptoms (Fdm.) | .82 | .81 |
| Social Adjustment | .89 | .93 |
| A. Social Standards | .84 | .76 |
| B. Social Skills | .86 | .70 |
| C. Anti-Social Tendencies (Fdm.) | .84 | .75 |
| D. Family Relations | .91 | .91 |
| E. School Relations | .73 | .84 |
| F. Community Relations | .78 | .77 |
| Total Adjustment | .93 | .95 |

| California Test of Personality Validity: | Content & Construct | Concurrent & Predictive |
|---|---|---|
| Personal Adjustment | Very Good | Poor |
| A. Self-Reliance | Good | Poor |
| B. Sense of Personal Worth | Good | Poor |
| C. Sense of Personal Freedom | Good | Poor |
| D. Feeling of Belonging | Good | Poor |
| E. Withdrawing Tendencies | Good | Poor |
| F. Nervous Symptoms | Very Good | Poor |
| Social Adjustment | Very Good | Poor |
| A. Social Standards | Good | Poor |
| B. Social Skills | Good | Poor |
| C. Anti-Social Tendencies | Good | Poor |
| D. Family Relations | Good | Poor |
| E. School Relations | Good | Poor |
| F. Community Relations | Good | Poor |
| Total Adjustment | | |

California Test of Personality (Critique)

See Also:

Norman, R.D. and Daley, M.F. "The Comparative Personality Adjustment of Superior and Inferior Readers." Journal of Educational Psychology, 1959, 50(1), 31-36.

## Columbia Mental Maturity Scale (1 Form)

**Purpose:** To estimate the intellectual ability of children

**Type of Test:** Individual intelligence

**Range:** 3-10 years (mental age)

**Time for Administration:** 15-20 minutes

**Cost:** $34.00 per kit, including 100 record blanks

**Subtests:** None

**Scores:** Single scaled IQ; Mean = 100, S.D. = 15

### Nature of Test:

The Scale includes 100 items, each printed on a card 6 by 19 inches, bearing from 3 to 5 drawings printed in various colors on white stock, or in black on colored stock. The cards are graded in difficulty from the simple to the more complex.

The child is required to select, after he has looked at a card for about 20 seconds, the drawing that is different from the others. Responses are recorded on an 8 1/2" x 11" Individual Record Blank. The child may reply by pointing or nodding his head at one of the drawings on each of these large, easily handled cards. The maximum possible score is 100.

### Columbia Mental Maturity - Means, Standard Deviations

Columbia Mental Maturity Scale IQ's for Separate Age Groups in Standardization Group

| Age | Number of Cases | Columbia IQ Mean | SD |
|---|---|---|---|
| 3-0 to 3-11 | 89 | 113.7 | 20.0 |
| 4-0 to 4-11 | 139 | 106.1 | 21.0 |
| 5-0 to 5-11 | 113 | 102.8 | 17.7 |
| 6-0 to 6-11 | 124 | 111.3 | 23.1 |
| 7-0 to 7-11 | 125 | 106.4 | 22.1 |
| 8-0 to 8-11 | 127 | 99.3 | 23.2 |
| 9-0 to 9-11 | 82 | 104.6 | 26.0 |
| 10-0 to 10-11 | 79 | 99.9 | 26.6 |
| 11-0 to 11-11 | 51 | 102.0 | 25.9 |
| 12-0 to 12-11 | 28 | 100.4 | 25.0 |

### Columbia Mental Maturity Scale for Selected Age Groups

**Reliability:**

| Age | r |
|---|---|
| 4 | .89 |
| 5 | .91 |
| 9 | .91 |
| 10 | .92 |

**Validity:** .45-.55

Deutsch SES Index

The index of cultural deprivation (C.D.) includes the following three components:

1. The father's occupation on a 4-point scale (or the mother's occupation if no father is in the home).

1 point - unskilled or unemployed
2 points - semiskilled
3 points - skilled
4 points - professional

2. Number of years of education completed by the parents (an average of the two, or the mother's education only if no father is in the home).

3. Density in the home is defined as the number of rooms divided by the number of people living in the home.

> In the number of rooms are included the kitchen and the bathroom. A shared bathroom is counted as half a room.

Occupation and education are given a full weight, but density is given a 1/2 weight. Each component is divided by its standard deviation to equate the variability of all components. In other words, the C.D. rating is the sum of three z scores, except that density is given a 1/2 weight.

The above gives a definitional formula. In actuality, the following computation formula is used:

$$1/2\ (\text{Education}) + 2\ (\text{Occupation}) + 2\ \frac{\text{Rooms}}{\text{People}}$$

Reflected in the computational formula are the following approximate standard deviations from the original Perry School population: 2 for education, 1/2 for occupation, and 1/4 for density.*

An example from an actual case may best clarify the computational formula. If pertinent data are:

Father's occupation: Unskilled factory work (1 point)
Mother's education: 11 years
Father's education: 7 years
Number of rooms: 6
Number of people living in home: 9

$$\text{C.D.} = 1/2(9) + 2(1) + 2\left(\frac{6}{9}\right) = 8.3$$

---

*A score of 10 or below makes the family a low SES family and thus "culturally deprived."

Illinois Test of Psycholinguistic Abilities (ITPA) (1 Form)

Purpose: To identify "psycholinguistic abilities and disabilities in children."

Type of Test: Individual psycholinguistic

Range: 2.5-9 years

Time for Administration: 45-60 minutes

Cost: $32.00 per kit; $3.75 per 25 records; $3.50 per manual

Subtests:
- Auditory Decoding
- Auditory Vocal Association
- Auditory Vocal Automatic
- Auditory Vocal Sequencing
- Vocal Encoding
- Visual Decoding
- Visual Motor Association
- Visual Motor Sequencing
- Motor Encoding

Scores: Profile of nine language skills (from subtests) and (total language age)

Nature of Test:

The psycholinguistic model on which the ITPA is based attempts to relate those functions whereby the intentions of one individual are transmitted (verbally or nonverbally) to another individual, and reciprocally, functions whereby the environment or the intentions of another individual are received and interpreted.

The adoption of this theoretical model for the battery served two purposes: (a) it was a parsimonious device by which the essential features of communication were delineated so that their relationships were specified; (b) it provided a framework within which to observe and evaluate a child, making it possible to verify and elaborate on test results and to suggest remedial measures.

The present model postulates three dimensions of cognitive abilities:

1. Channels of Communication. These are the routes through which the content of communication flows. Included here are the modalities through which sense impressions are received and the forms of expression through which a response is made. The channels may include various combinations of sensory input and response output. The major modes of input are auditory and visual; those of putput are vocal and motor. Due to practical limitation, the ITPA incorporates only the auditory-vocal and the visual-motor channels. These channels were selected as being most relevant for the developmental level of subjects in the test's age range.
2. Psycholinguistic Processes. In analyzing behavior which occurs in the acquisition and use of language, three main processes are considered: (a) the receptive process, that is, that ability necessary to recognize and/or understand what is seen or heard; (b) the expressive process, that is, those skills necessary to express ideas or to respond either vocally or by gesture or movement; (c) an organizing process which involves the internal manipulation of percepts, concepts, and linguistic symbols. It is a central mediating process elicited by the receptive process and preceding the expressive process.

3. Level of Organization. The degree to which habits of communication are organized within the individual determines the level of functioning. Two levels are postulated in the clinical model of the ITPA: (a) the representational level, which requires the more complex mediating process of utilizing symbols which carry the meaning of an object; (b) the automatic level in which the individual's habits of functioning are less voluntary but highly organized and integrated. The automatic chain of responses of the latter level is involved in such activities as visual and auditory closure, speed of perception, ability to reproduce a sequence seen or heard, rote learning, synthesizing isolated sounds into a word, and utilizing the redundancies of experience.

Psycholinguistic Age Norms for the 12 ITPA Subtests

| Raw Score | Aud. Recept. (Decoding) | Visual Recept. (Decoding) | Aud.-Voc. Assoc. | Vis.-Motor Assoc. | Verbal (Voc) Express. Encoding | Manual (Motor Express. Encoding) |
|---|---|---|---|---|---|---|
| 0 | 2-1 | | 2-4 | 2-4 | 2-0 | 2-0 |
| 1 | 2-2 | 2-1 | 2-6 | 2-6 | 2-2 | 2-2 |
| 2 | 2-4 | 2-4 | 2-8 | 2-8 | 2-5 | 2-4 |
| 3 | 2-6 | 2-7 | 2-10 | 2-10 | 2-7 | 2-6 |
| 4 | 2-8 | 2-10 | 3-0 | 3-0 | 2-9 | 2-8 |
| 5 | 2-10 | 3-1 | 3-3 | 3-2 | 2-10 | 2-10 |
| 6 | 2-11 | 3-4 | 3-5 | 3-4 | 3-1 | 3-0 |
| 7 | 3-1 | 3-7 | 3-7 | 3-6 | 3-4 | 3-2 |
| 8 | 3-3 | 3-10 | 3-9 | 3-8 | 3-8 | 3-4 |
| 9 | 3-5 | 4-1 | 3-11 | 3-10 | 3-10 | 3-6 |
| 10 | 3-6 | 4-4 | 4-1 | 4-0 | 4-0 | 3-8 |
| 11 | 3-8 | 4-7 | 4-3 | 4-2 | 4-3 | 3-10 |
| 12 | 3-10 | 4-10 | 4-5 | 4-4 | 4-6 | 4-0 |
| 13 | 4-0 | 5-0 | 4-7 | 4-6 | 4-8 | 4-4 |
| 14 | 4-1 | 5-2 | 4-9 | 4-8 | 4-10 | 4-6 |
| 15 | 4-3 | 5-5 | 4-11 | 4-10 | 4-0 | 4-10 |
| 16 | 4-5 | 5-7 | 5-1 | 5-0 | 5-2 | 5-0 |
| 17 | 4-7 | 5-10 | 5-3 | 5-3 | 5-4 | 5-3 |
| 18 | 4-10 | 6-0 | 5-5 | 5-6 | 5-6 | 5-6 |
| 19 | 5-0 | 6-2 | 5-7 | 5-9 | 5-8 | 5-10 |
| 20 | 5-2 | 6-5 | 5-9 | 6-0 | 5-10 | 6-1 |
| 21 | 5-4 | 6-7 | 6-0 | 6-3 | 6-0 | 6-5 |
| 22 | 5-6 | 6-10 | 6-2 | 6-6 | 6-2 | 6-9 |
| 23 | 5-8 | 7-1 | 6-6 | 6-10 | 6-4 | 7-2 |
| 24 | 5-10 | 7-4 | 6-9 | 7-2 | 6-6 | 7-6 |
| 25 | 6-0 | 7-9 | 7-0 | 7-7 | 6-8 | 7-11 |
| 26 | 6-3 | 8-4 | 7-3 | 8-0 | 6-10 | 8-4 |
| 27 | 6-5 | 8-10 | 7-8 | 8-5 | 7-0 | 8-8 |
| 28 | 6-8 | 9-3 | 7-11 | 8-11 | 7-3 | 9-2 |
| 29 | 6-10 | 9-10 | 8-3 | 9-4 | 7-5 | 9-10 |
| 30 | 7-0 | 10-4 | 8-6 | 9-10 | 7-8 | 10-4 |
| 31 | 7-3 | 10-10 | 8-10 | 10-3 | 7-10 | |
| 32 | 7-6 | | 9-2 | | 8-1 | |
| 33 | 7-9 | | 9-5 | | 8-4 | |

(continued)

Psycholinguistic Age Norms for the 12 ITPA Subtests

| Raw Score | Aud. Recept. (Decoding) | Visual Recept. (Decoding) | Aud.-Voc. Assoc. | Vis.-Motor Assoc. | Verbal(Voc) Express. Encoding | Manual(Motor Express. Encoding) |
|---|---|---|---|---|---|---|
| 34 | 8-0 | | 9-9 | | | |
| 35 | 8-4 | | 10-1 | | | |
| 36 | 8-7 | | 10-6 | | | |
| 37 | 8-10 | | 10-11 | | | |
| 38 | 9-2 | | | | | |
| 39 | 9-6 | | | | | |
| 40 | 9-10 | | | | | |
| 41 | 10-2 | | | | | |

Psycholinguistic Age Norms for the 12 ITPA Subtests (continued)

| Raw Score | Vis.-Motor Assoc. | Verbal(Voc) Express. Encoding) | Manual(Motor Express. Encoding) | Grammatic Clos. (Aud. Vocal Automatic) | Visual Clos. | Aud. Sequen. Memory(Aud.Voc. Sequencing) |
|---|---|---|---|---|---|---|
| 0 | 2-4 | 2-0 | | | | |
| 1 | 2-6 | 2-2 | 2-0 | 2-2 | | |
| 2 | 2-8 | 2-5 | 2-2 | 2-6 | | |
| 3 | 2-10 | 2-7 | 2-4 | 2-10 | 2-2 | 2-0 |
| 4 | 3-0 | 2-9 | 2-6 | 3-3 | 2-6 | 2-2 |
| 5 | 3-2 | 2-10 | 2-8 | 3-7 | 2-11 | 2-5 |
| 6 | 3-4 | 3-1 | 2-10 | 3-10 | 3-3 | 2-7 |
| 7 | 3-6 | 3-4 | 3-0 | 4-2 | 3-6 | 2-10 |
| 8 | 3-8 | 3-8 | 3-2 | 4-5 | 3-10 | 3-0 |
| 9 | 3-10 | 3-10 | 3-4 | 4-8 | 4-0 | 3-2 |
| 10 | 4-0 | 4-0 | 3-6 | 4-10 | 4-4 | 3-5 |
| 11 | 4-2 | 4-3 | 3-8 | 5-0 | 4-6 | 3-7 |
| 12 | 4-4 | 4-6 | 3-10 | 5-2 | 4-8 | 3-10 |
| 13 | 4-6 | 4-8 | 4-0 | 5-4 | 4-10 | 4-0 |
| 14 | 4-8 | 4-10 | 4-4 | 5-6 | 5-0 | 4-2 |
| 15 | 4-10 | 4-0 | 4-6 | 5-8 | 5-3 | 4-5 |
| 16 | 5-0 | 5-2 | 4-10 | 5-10 | 5-5 | 4-8 |
| 17 | 5-3 | 5-4 | 5-0 | 6-0 | 5-8 | 4-10 |
| 18 | 5-6 | 5-6 | 5-3 | 6-2 | 5-10 | 5-0 |
| 19 | 5-9 | 5-8 | 5-6 | 6-5 | 6-1 | 5-3 |
| 20 | 6-0 | 5-10 | 5-10 | 6-8 | 6-4 | 5-6 |
| 21 | 6-3 | 6-0 | 6-1 | 7-0 | 6-6 | 5-8 |
| 22 | 6-6 | 6-2 | 6-5 | 7-3 | 6-9 | 6-0 |
| 23 | 6-10 | 6-4 | 6-9 | 7-7 | 7-0 | 6-3 |
| 24 | 7-2 | 6-6 | 7-2 | 7-11 | 7-2 | 6-6 |
| 25 | 7-7 | 6-8 | 7-6 | 8-2 | 7-6 | 6-10 |
| 26 | 8-0 | 6-10 | 7-11 | 8-6 | 7-9 | 7-2 |
| 27 | 8-5 | 7-0 | 8-4 | 8-10 | 8-0 | 7-7 |
| 28 | 8-11 | 7-3 | 8-8 | 9-2 | 8-3 | 7-11 |
| 29 | 9-4 | 7-5 | 9-2 | 9-8 | 8-6 | 8-4 |
| 30 | 9-10 | 7-8 | 9-10 | 10-4 | 8-9 | 8-8 |
| 31 | 10-3 | 7-10 | 10-4 | | 9-0 | 9-2 |
| 32 | | 8-1 | | | 9-4 | 9-10 |
| 33 | | 8-4 | | | 9-10 | 10-3 |
| 34 | | 8-7 | | | 10-6 | |
| 35 | | 8-10 | | | | |
| 36 | | 9-1 | | | | |
| 37 | | 9-4 | | | | |
| 38 | | 9-10 | | | | |
| 39 | | 10-2 | | | | |
| 40 | | 10-5 | | | | |
| 41 | | 10-11 | | | | |

Psycholinguistic Age Norms for the 12 ITPA Subtests (continued)

| Raw Score | Vis. Sequen. Memory(Vis.Motor Sequencing) | Aud. Clos. | Sound Blending |
|---|---|---|---|
| | | Supplemental Tests of Grammatical Clos. | |
| 0 | | 2-0 | |
| 1 | 2-1 | 2-4 | |
| 2 | 2-4 | 2-7 | 2-4 |
| 3 | 2-7 | 2-10 | 2-10 |
| 4 | 2-10 | 3-1 | 3-4 |
| 5 | 3-1 | 3-4 | 3-8 |
| 6 | 3-4 | 3-7 | 4-2 |
| 7 | 3-7 | 3-10 | 4-6 |
| 8 | 3-10 | 4-0 | 5-0 |
| 9 | 4-1 | 4-2 | 5-3 |
| 10 | 4-4 | 4-4 | 5-6 |
| 11 | 4-7 | 4-6 | 5-10 |
| 12 | 4-10 | 4-8 | 6-2 |
| 13 | 5-1 | 4-11 | 6-5 |
| 14 | 5-4 | 5-1 | 6-8 |
| 15 | 5-7 | 5-3 | 6-11 |
| 16 | 5-10 | 5-6 | 7-1 |
| 17 | 6-2 | 5-9 | 7-4 |
| 18 | 6-6 | 6-1 | 7-7 |
| 19 | 6-10 | 6-5 | 7-10 |
| 20 | 7-3 | 6-9 | 8-2 |
| 21 | 7-10 | 7-3 | 8-7 |
| 22 | 8-4 | 7-11 | |
| 23 | 8-9 | 8-5 | |
| 24 | 10-5 | 9-4 | |
| 25 | | 10-0 | |
| 26 | | | |
| 27 | | | |
| 28 | | | |
| 29 | | | |
| 30 | | | |
| 31 | | | |
| 32 | | | |
| 33 | | | |

Illinois Test of Psycholinguistic Abilities (ITPA)

| Reliability | Median Reliability | Range |
|---|---|---|
| Auditory-Vocal Automatic | .78 | .62-.90 |
| Visual Decoding | .76 | .45-.90 |
| Motor Encoding | .76 | .60-.86 |
| Auditory-Vocal Association | .72 | .51-.83 |
| Visual Motor Sequential | .70 | .39-.84 |
| Vocal Encoding | .72 | .48-.84 |
| Auditory-Vocal Sequencing | .80 | .72-.89 |
| Visual Motor Association | .76 | .55-.89 |
| Auditory Decoding | .80 | .58-.91 |
| **Validity** | | |
| Concurrent | .15 | .03-.65 |
| Predictive | .23 | -.19-.53 |

Illinois Test of Psycholinguistic Abilities (Critiques)

See Also:

Hirshoren, A. "A Comparison of the Predictive Validity of the Revised Stanford-Binet Intelligence Scale and the Illinois Test of Psycholinguistic Abilities," 1969, 35(3), 517-521.

Kirk, S.A. and McCarthy, J.J. "The Illinois Test of Psycholinguistic Abilities: An Approach to Differential Diagnosis." American Journal of Mental Deficiency, 1961, 66(3), 399-412.

Mueller, M.W. "Validity of Six Tests of Ability with Educable Mental Retardates." Journal of School Psychology, 1968, 6(2), 136-146.

Weener, P., Barritt, L.S., and Semmel, M.I. "A Critical Evaluation of the Illinois Test of Psycholinguistic Abilities." Exceptional Children, 1967, 33(2), 373-380.

Leiter International Performance Scale (1 Form)

Purpose: To assess intelligence (has no spoken instructions)

Type of Test: Individual intelligence

Range: 2-18 years

Time for Administration: Untimed (usually 1 hour)

Cost: $350.00 per set (includes test materials, 100 record blanks, carry case and manual)

Subtests:
- Matching
- Copying
- Picture Completion
- Number Estimation
- Analogies
- Recognition of Age
- Classification of Animals According to Habitat
- Series Completion
- Spatial Relations
- Footprint Recognition
- Similarities
- Memory for a Series

Scores: Single IQ, scaled

Mean = 100 Standard Deviation = 15

Nature of Test:

The Leiter Scale is designed to cover a wide range of functions, similar to those found in verbal scales. A distinctive feature is the absence of spoken instructions, thereby requiring no language. It is appropriate especially for handicapped children and those children for whom culture bias might otherwise be a factor.

Reliability: .90

Validity: Good

Metropolitan Achievement Test (2 Forms except for Primary I(1) )

Purpose: To measure "achievement in the important skill and content areas of the elementary and junior high school curriculum"

Type of Test: Group achievement

Range: Grades 1.5-12

Time for Administration: Varies from 85-316 minutes

Subject Tests, Number of Items and Working Time

| Subtest | Primary I | | Primary II | | Elemen. | | Intermed. | | Advanced | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Items | Time | Items | Time | Items | Time | Items | Time | Items | Time |
| Word Knowledge | 35 | 15 | 37 | 18 | 50 | 15 | 55 | 14 | 55 | 14 |
| Word Discrimination | 35 | 12 | 35 | 12 | 36 | 12 | | | | |
| Reading | 45 | 35 | 51 | 35 | 44 | 22 | 44 | 25 | 45 | 25 |
| Spelling | | | 30 | 10 | 40 | 20 | 55 | 17 | 55 | 17 |
| Arith:Concepts & Skills | 63 | 23 | | | | | | | | |
| Prob. Solving & Concepts | | | 42 | 15 | 35 | 26 | 48 | 35 | 48 | 35 |
| Computation | | | 30 | 10 | 47 | 30 | 48 | 38 | 45 | 35 |
| Language:Usage | | | | | 24 | 10 | 35 | 13 | 24 | 8 |
| Punctuation & Capitalization | | | | | 36 | 12 | 36 | 9 | 38 | 7 |
| Parts of Speech-Grammar | | | | | | | 10 | 4 | 20 | 10 |
| Kinds of Sent. | | | | | | | | | 10 | 4 |
| Language Study Skills | | | | | | | 28 | 16 | 28 | 16 |
| Social Studies Information | | | | | | | 60 | 20 | 60 | 20 |
| Social Studies Study Skills | | | | | | | 29 | 26 | 39 | 36 |
| Science | | | | | | | 55 | 20 | 55 | 20 |

Cost: $6.25 per 35 tests, $1.20 per manual, $2.00 per 35 cummulative record cards.

Scores: Raw; one for each subtest.

Nature of Test:

Metropolitan Achievement Tests comprise a coordinated series of measures of achievement. At each level Metropolitan includes tests covering the most important knowledge or skill areas in the grade or grades for which that level is intended. The tests in each battery, the number of questions in each test and the working time required are summarized in the chart above. The various subtests yield comparable results so that the teacher may readily infer particular strengths and weaknesses of a pupil or class.

Subtests:

Primary I (Grade 1.5-1.9)

| | |
|---|---|
| Word Knowledge | Reading |
| Word Discrimination | Arith:Concepts & Skills |

Primary II (Grade 2)

| | |
|---|---|
| Word Knowledge | Reading |
| Word Discrimination | Problem Solving & Concepts |
| Spelling | Computation |

Elementary (Grades 3-4)

| | |
|---|---|
| Word Knowledge | Reading |
| Word Discrimination | Problem Solving & Concepts |
| Spelling | Computation |
| Language Usage | Punct. & Capitalization |

Intermediate (Grades 5-6)

| | |
|---|---|
| Word Knowledge | Reading |
| Spelling | Problem Solving & Concepts |
| Language Usage | Computation |
| Parts of Speech-Grammar | Punct. & Capitalization |
| Language Study Skills | Soc. Studies Information |
| Social Studies Study Skills | Science |

Advanced (Grades 7-9)

| | |
|---|---|
| Word Knowledge | Reading |
| Spelling | Problem Solving & Concepts |
| Language Usage | Computation |
| Parts of Speech-Grammar | Punct. & Capitalization |
| Kinds of Sentences | Language Study Skills |
| Social Studies Information | Social Studies Skills |
| Science | |

High School Battery (Grades 9-12)

| | |
|---|---|
| Reading | Social Studies Information |
| Spelling | Math. Computation & Concepts |
| Language | Math. Analy. & Prob. Solving |
| Language Study Skills | Scient.Concepts & Understand. |
| Social Studies Study Skills | Scientific Information |
| Social Studies Vocabulary | |

Subtests:

| Grade 1 (Primary I) | Means | SD |
|---|---|---|
| Reading:Word Pictures (Knowledge) | 24.81 | 10.89 |
| Reading:Word Recognition (Reading) | 17.88 | 5.67 |
| Reading:Word Meaning (Discrimination) | 14.84 | 8.88 |
| Numbers (Concepts & Skills) | 47.52 | 12.82 |
| **Grade 2 (Primary II)** | | |
| Reading | 33.36 | 11.83 |
| Word Meaning (Word Discrimination) | 48.82 | 22.48 |
| Arith. Fundamentals (Computation) | 44.88 | 12.77 |
| Arith. Problems (Prob. Solving & Concepts) | 11.78 | 8.40 |
| Spelling | 15.68 | 6.28 |

| Subtests: | Means | SD |
|---|---|---|
| **Elementary (Grade 4.1)** | | |
| Word Knowledge | 33.0 | 11.55 |
| Reading | 26.4 | 9.37 |
| Spelling | 25.6 | 11.36 |
| Math. Computation | 24.2 | 7.40 |
| Math. Concepts | 23.8 | 8.07 |
| **Intermediate (Grade 5.1)** | | |
| Word Knowledge | 26.0 | 10.09 |
| Reading | 22.5 | 9.73 |
| Spelling | 22.8 | 9.12 |
| Math. Computation | 17.1 | 6.69 |
| Math. Concepts | 18.3 | 7.60 |
| Science | 38.1 | 15.13 |
| Social Studies | 42.4 | 17.22 |
| **Intermediate (Grade 6.1)** | | |
| Word Knowledge | 30.8 | 10.37 |
| Reading | 26.2 | 10.46 |
| Spelling | 28.6 | 9.83 |
| Math. Computation | 23.3 | 7.83 |
| Math. Concepts | 22.7 | 8.11 |
| Science | 49.6 | 15.70 |
| Social Studies | 54.1 | 19.12 |

**Metropolitan Achievement Tests**

Reliability:

**Primary 1 Battery (Last half of Grade 1)**

| Test: | Range | Mdn. |
|---|---|---|
| Word Knowledge | .86-.91 | .90 |
| Word Discrimination | .81-.83 | .87 |
| Reading | .88-.94 | .92 |
| Arithmetic | | |
| Concepts (items 1-39) | .81-.89 | .85 |
| Skills (items 40-63) | .94-.95 | .95 |
| Total (items 1-63) | .91-.95 | .93 |

**Primary II Battery (Grade 2)**

| Test: | Range | Mdn. |
|---|---|---|
| Word Knowledge | .86-.96 | .93 |
| Word Discrimination | .86-.92 | .88 |
| Reading | .94-.95 | .94 |
| Spelling | .92-.95 | .93 |
| Arithmetic: | | |
| Part A Concepts & Problem | .80-.87 | .86 |
| Solving | .74-.88 | .80 |
| Part B Computation | .85-.92 | .91 |
| Total (Parts A & B) | | |

## Metropolitan Achievement Tests

Reliability:

### Elementary Battery (Grades 3-4)

| Test: | Range | Mdn. |
|---|---|---|
| Word Knowledge | .79-.95 | .93 |
| Word Discrimination | .88-.93 | .92 |
| Reading | .86-.91 | .90 |
| Spelling | .94-.96 | .95 |
| Language: | | |
| Part A Usage | .74-.84 | .80 |
| Part B Punctuation & Capitalization | .61-.81 | .78 |
| Total (Parts A & B) | .79-.87 | .84 |
| Arith. Computation | .91-.93 | .92 |
| Arith. Problem Solving & Concepts | .86-.91 | .88 |

### Intermediate Battery (Grades 5-6)

| Test: | Range | Mdn. |
|---|---|---|
| Word Knowledge | .88-.95 | .94 |
| Reading | .89-.92 | .90 |
| Spelling | .91-.96 | .92 |
| Language: | | |
| Part A Usage | .78-.84 | .81 |
| Part B Parts of Speech | .64-.77 | .72 |
| Part C Punctuation & Capitalization | .80-.88 | .83 |
| Total (Parts A-C) | .87-.91 | .89 |
| Language Study Skills | .76-.85 | .79 |
| Arith. Computation | .82-.94 | .88 |
| Arith. Problem Solving & Concepts | .90-.95 | .92 |
| Social Studies Information | .86-.87 | .87 |
| Social Studies Study Skills | .64-.77 | .73 |
| Science | .87-.90 | .89 |

### Advanced Battery (Grades 7-9)

| Test: | Range |
|---|---|
| Word Knowledge | .91-.94 |
| Reading | .88-.92 |
| Spelling | .90-.93 |
| Language: | |
| Part A Usage | .47-.81 |
| Part B Punctuation & Capitalization | .63-.77 |
| Part C Kinds of Sentences | .53-.67 |
| Part D Parts of Speech | .81-.86 |
| Total (Parts A-D) | .88-.90 |
| Language Study Skills | .79-.84 |
| Arith. Computation | .87-.95 |
| Arith. Problem Solving & Concepts | .89-.94 |
| Social Studies Information | .86-.91 |
| Social Studies Study Skills | .82-.86 |
| Science | .79-.88 |

Metropolitan Achievement Tests

| Validity: | Content & Construct | Concurrent & Predictive |
|---|---|---|
| | Generally Very Good | Not Reported |

Metropolitan Achievement Tests (Critiques)

See Also:

Eagle, N. and Harris, A.S. "Interaction of Race and Test on Reading Performance Scores." Journal of Educational Measurement, 1966, 6(3), 131-135.

Metropolitan Readiness Tests (2 Forms)

Purpose: To measure the traits and achievements of school beginners that contribute to their readiness for first grade instruction.

Type of Test: Group Readiness

Range: End of kindergarten or first grade

Time for Administration: 60 minutes

Cost: $2.35 per 25 tests

Subtests:
- Word Meaning
- Information
- Numbers
- Drawing a Man (optional)
- Sentences
- Matching
- Copying

Scores: Raw: one for each subtest, and total score (sum of subtest scores)

Nature of Test:

Metropolitan Readiness Tests were designed for testing pupils at the end of the kindergarten year or the beginning of the first grade. These tests provide a quick, convenient, and dependable basis for early classification of pupils, thus helping teachers manage the instructional effort more efficiently.

Six tests are included in Metropolitan Readiness Tests, as follows:

Test 1: Word Meaning, 16-item picture vocabulary test. The pupil selects from three pictures the one that illustrates the word the examiner names.

Test 2: Listening, a 16-item test of ability to comprehend phrases and sentences instead of individual words. The pupil selects from three pictures the one which portrays a situation or even the examiner describes briefly.

Test 3: Matching, a 14-item test of visual perception involving the recognition of similarities. The pupil marks one of three pictures which matches a given picture.

Test 4: Alphabet, a 16-item test of ability to recognize lower-case letters of the alphabet. The pupil chooses a letter named from among four alternatives.

Test 5: Numbers, a 26-item test which measures a combination of visual perception and motor control.

A seventh, optional, test, Draw-a-Man, provides an index of general intellectual maturity. This is an adaptation of one drawing from the Goodenough-Harris Drawing Test.

Metropolitan Readiness Tests

Means and Standard Deviations, Grade 1

| Test: | 1st Testing Mean | SD | 2nd Testing Mean | SD |
|---|---|---|---|---|
| Word Meaning | 15.61 | 2.51 | 15.92 | 2.26 |
| Sentences | 10.52 | 2.39 | 10.86 | 2.27 |
| Information | 12.02 | 2.06 | 12.39 | 1.89 |
| Matching | 13.19 | 4.23 | 14.19 | 3.75 |
| Numbers | 13.50 | 4.78 | 14.38 | 4.85 |
| Copying | 5.26 | 2.76 | 5.31 | 2.75 |
| Tests 1-4 | 51.07 | 8.83 | 53.20 | 8.13 |
| Tests 1-6 | 69.71 | 13.92 | 72.96 | 13.12 |

Metropolitan Readiness Tests

| Reliability: | Test | Reliability Coefficient |
|---|---|---|
| | Word Meaning | .583 |
| | Sentences | .535 |
| | Information | .586 |
| | Matching | .773 |
| | Numbers | .839 |
| | Copying | .762 |
| | Tests 1-4 | .828 |
| | Tests 1-6 | .890 |

| Validity: | Test | Average Reading | Numbers |
|---|---|---|---|
| | Test 1 Word Meaning | .409 | .441 |
| | Test 2 Sentences | .418 | .424 |
| | Test 3 Information | .341 | .400 |
| | Test 4 Matching | .486 | .541 |
| | Test 5 Numbers | .520 | .634 |
| | Test 6 Copying | .341 | .402 |
| | Tests 1-4 Reading Readiness | .475 | .530 |
| | Tests 1-6 Total Readiness | .534 | .615 |

Metropolitan Readiness Tests (Critiques)

See Also:

Olson, A.V. and Johnson, C.I. "Structure and Predictive Validity of the Frostig Developmental Test of Visual Perception in Grades One and Three." The Journal of Special Education, 1970, 4(1), 49-52.

Peabody Picture Vocabulary Test (2 Forms)

Purpose: To provide an estimate of a subject's verbal intelligence through measuring his hearing vocabulary

Type of Test: Individual intelligence

Range: 2.5-18.5 years

Time for Administration: Untimed (usually 10-15 minutes)

Cost: $3.00 per 50 copies

Subtests: None

Scores: Total IQ

Standardized Mean = 100 Standard Deviation = 15

Nature of Test:

The test consists of 150 items, each of which consists of four pictures. The subject is asked to indicate which picture the examiner asks him to find. Besides being effective with average subjects, it has special value with certain other groups. Since no reading is required, it is especially fair for non-readers, and remedial cases. Because the answers may be indicated either orally or physically, it is advantageous for various handicapped persons. The illustrations are bold, clean line drawings without fine detail and figure-ground problems, thus making it more appropriate for visually-impaired and perceptually-handicapped individuals.

Peabody Picture Vocabulary Test

| Age Levels | Reliability Coefficients For Raw Scores | Standard Errors For I.Q. Scores |
|---|---|---|
| 18-0 | 0.84 | 6.00 |
| 17-0 | 0.84 | 6.00 |
| 16-0 | 0.80 | 6.70 |
| 15-0 | 0.83 | 6.18 |
| 14-0 | 0.81 | 6.54 |
| 13-0 | 0.70 | 8.22 |
| 12-0 | 0.78 | 7.04 |
| 11-0 | 0.81 | 6.54 |
| 10-0 | 0.77 | 7.20 |
| 9-0 | 0.74 | 7.65 |
| 8-0 | 0.79 | 6.87 |
| 7-0 | 0.74 | 7.65 |
| 6-0 | 0.67 | 8.61 |
| 5-0 | 0.73 | 7.80 |
| 4-6 | 0.72 | 7.94 |
| 4-0 | 0.77 | 7.20 |
| 3-6 | 0.81 | 6.54 |
| 3-0 | 0.75 | 7.50 |
| 2-6 | 0.76 | 7.35 |

Validity: Congruent - .71 (Stanford-Binet)
Congruent - .61 (Wechsler Scales)
Concurrent - .50+
Predictive - .38

Peabody Picture Vocabulary Test (Critiques)

Gage, G.E. and Naumann, T.F. "Correlation of the Peabody Picture Vocabulary Test and the Wechsler Intelligence Scale for Children." The Journal of Educational Research, 1965, 58(10), 466-468.

Mueller, M.W. "Validity of Six Tests of Ability with Educable Mental Retardates." Journal of School Psychology, 1968, 6(2), 136-146.

Nelson, C.M. and Hudson, F.G. "Predicting the Reading Achievement of Junior High School EMR Children." American Journal of Mental Deficiency, 1969, 74(3), 415-420.

Pool, D.A. and Brown, R. "The Peabody Picture Vocabulary Test as a Measure of General Adult Intelligence." Journal of Consulting and Clinical Psychology, 1970, 34(1), 8-11.

Raven Progressive Matrices (1 Form)

Purpose: To test observation and clear thinking

Type of Test: Individual intelligence

Range: 5-11 years (may also be applicable for younger children)

Time for Administration: 30 minutes

Cost: $14.50 per 25 tests, $2.00 per manual

Subtests: None

Scores: Single score
Standardized Mean = 100, S.D. = 15

Nature of Test:

The Progressive Matrices is a test of a person's capacity at the time of the test to apprehend meaningless figures presented for his observation, see the relations between them, conceive the nature of the figure completing each system of relations presented, and, by so doing, develop a systematic method of reasoning. The scale consists of 60 problems divided into five sets of twelve. In each set the first problem is as nearly as possible self-evident. The problems which follow become progressively more difficult. The scale is intended to cover the whole range of intellectual development.

Name of Test: Raven Progressive Matrices

Comparison Between Test Results for 25 Normal and 29 Disturbed Children, 6 1/2- 12 1/2 Years of Age, Given the Matrices and the Vocabulary Scale Three Times at Intervals of Three Months

| Tests given at intervals of 3 months | | Normal School Children | | | Emotionally Disturbed Children | | |
|---|---|---|---|---|---|---|---|
| | | 1st | 2nd | 3rd | 1st | 2nd | 3rd |
| TEST SCORES | P.M. Mean Score | 24.9 | 27.2 | 28.9 | 20.5 | 21.9 | 23.4 |
| | Standard Dev. | 5.8 | 6.3 | 7.1 | 6.1 | 7.2 | 7.4 |

Raven Progressive Matrices

Reliability: .88
Validity: Content and Construct - Good
Concurrent and Predictive - Fair

Raven Progressive Matrices (Critiques)

See Also:

Knief, L.M. and Stroud, J.B. "Intercorrelations Among Various Intelligence, Achievement, and Social Class Scores." Journal of Educational Psychology, 1959, 50(3), 117-120.

Mueller, M.W. "Validity of Six Tests of Ability with Educable Mental Retardates." Journal of School Psychology, 1968, VI(2), 136-146.

Stanford Achievement Tests (3 Forms)

Purpose: To test the educational achievements that are commonly expected of students in a modern comprehensive school

Type of Test: Individual or group achievement

Range: Grades 1-12

Time for Administration: Varies from 127-440 minutes (4-7 sessions)

Cost: $5.65-$9.75 per 35 tests, $.60-$1.60 per set of keys

Subtests:

Primary I Battery (Grades 1.5-2.5)

| | |
|---|---|
| Word Reading (Word Meaning) | Spelling |
| Paragraph Meaning | Word Study Skills |
| Vocabulary | Arithmetic (Concepts) |

Primary II Battery (Grades 2.5-3.9)

| | |
|---|---|
| Word Meaning | Spelling |
| Paragraph Meaning | Word Study Skills |
| Science & Social Studies Concepts | Language |
| Arithmetic Concepts | Arithmetic Computation |

Intermediate I Battery (Grades 4.0-5.5)

| | |
|---|---|
| Word Meaning | Social Studies |
| Paragraph Meaning | Arithmetic Computation |
| Spelling | Arithmetic Concepts |
| Word Study Skills | Arithmetic Applications |
| Language | Science |

Intermediate II Battery (Grades 5.5-6.9)

| | |
|---|---|
| Word Meaning | Arithmetic Computation |
| Paragraph Meaning | Arithmetic Concepts |
| Spelling | Arithmetic Applications |
| Language | Social Studies |
| Science | |

Advanced Battery (Grades 7.0-9.9)

| | |
|---|---|
| Paragraph Meaning | Arithmetic Concepts |
| Spelling | Arithmetic Applications |
| Language | Social Studies |
| Arithmetic Computation | Science |

High School Battery (Grades 9-12)

| | |
|---|---|
| Reading | Mathematics |
| Spelling | Social Studies |
| English | Arts & Humanities |
| Numerical Competence | Business & Economics |
| Science | Technical Comprehension |

Scores: Raw; for each subtest

Stanford Achievement Tests

| | | | | | | |
|---|---|---|---|---|---|---|
| Subtest: | Word Reading | Grade 1 | Means | 16.1 | S.D. | 6.88 |
| | Paragraph Meaning | Grade 1 | Means | 14.1 | S.D. | 7.05 |
| | Vocabulary | Grade 1 | Means | 18.8 | S.D. | 6.69 |
| | Spelling | Grade 1 | Means | 7.7 | S.D. | 6.05 |
| | Word Study Skills | Grade 1 | Means | 31.7 | S.D. | 9.47 |
| | Arithmetic | Grade 1 | Means | 30.4 | S.D. | 14.43 |
| Subtest: | Word Meaning | Grade 2 | Means | 16.0 | S.D. | 6.3 |
| | Paragraph Meaning | Grade 2 | Means | 28.1 | S.D. | 12.35 |
| | Sci. & Social Studies Concepts | Grade 2 | Means | 17.6 | S.D. | 4.97 |
| | Spelling | Grade 2 | Means | 10.9 | S.D. | 7.24 |
| | Word Study Skills | Grade 2 | Means | 34.5 | S.D. | 12.90 |
| | Language | Grade 2 | Means | 35.3 | S.D. | 9.72 |
| | Arith. Computation | Grade 2 | Means | 18.4 | S.D. | 8.38 |
| | Arith. Concepts | Grade 2 | Means | 17.2 | S.D. | 7.42 |
| Subtest: | Word Meaning | Grade 3 | Means | 23.5 | S.D. | 6.68 |
| | Paragraph Meaning | Grade 3 | Means | 40.1 | S.D. | 11.33 |
| | Sci. & Social Studies Concepts | Grade 3 | Means | 22.1 | S.D. | 5.70 |
| | Spelling | Grade 3 | Means | 18.4 | S.D. | 7.73 |
| | Word Study Skills | Grade 3 | Means | 42.1 | S.D. | 13.23 |
| | Language | Grade 3 | Means | 46.4 | S.D. | 11.40 |
| | Arith. Computation | Grade 3 | Means | 32.4 | S.D. | 9.30 |
| | Arith. Concepts | Grade 3 | Means | 26.6 | S.D. | 9.58 |
| Subtest: | Word Meaning | Grade 4 | Means | 19.0 | S.D. | 7.94 |
| | Paragraph Meaning | Grade 4 | Means | 30.6 | S.D. | 11.22 |
| | Spelling | Grade 4 | Means | 28.6 | S.D. | 11.37 |
| | Word Study Skills | Grade 4 | Means | 37.9 | S.D. | 12.92 |
| | Language | Grade 4 | Means | 73.6 | S.D. | 17.15 |
| | Arith. Computation | Grade 4 | Means | 19.2 | S.D. | 7.07 |
| | Arith. Concepts | Grade 4 | Means | 15.4 | S.D. | 6.86 |
| | Arith. Applications | Grade 4 | Means | 16.0 | S.D. | 6.46 |
| | Social Studies | Grade 4 | Means | 25.8 | S.D. | 8.28 |
| | Science | Grade 4 | Means | 31.1 | S.D. | 11.16 |
| Subtest: | Word Meaning | Grade 5 | Means | 22.3 | S.D. | 8.76 |
| | Paragraph Meaning | Grade 5 | Means | 31.7 | S.D. | 11.96 |
| | Spelling | Grade 5 | Means | 26.8 | S.D. | 10.14 |
| | Language | Grade 5 | Means | 81.2 | S.D. | 18.72 |
| | Arith. Computation | Grade 5 | Means | 14.8 | S.D. | 5.18 |
| | Arith. Concepts | Grade 5 | Means | 12.1 | S.D. | 5.15 |
| | Arith. Applications | Grade 5 | Means | 16.6 | S.D. | 6.05 |
| | Social Studies | Grade 5 | Means | 37.3 | S.D. | 11.79 |
| | Science | Grade 5 | Means | 30.0 | S.D. | 9.53 |

(Continued)

Stanford Achievement Tests

| Subtest: | | | | | | |
|---|---|---|---|---|---|---|
| | Word Meaning | Grade 6 | Means | 27.6 | S.D. | 9.25 |
| | Paragraph Meaning | Grade 6 | Means | 38.4 | S.D. | 11.94 |
| | Spelling | Grade 6 | Means | 33.1 | S.D. | 11.26 |
| | Language | Grade 6 | Means | 90.5 | S.D. | 18.88 |
| | Arith. Computation | Grade 6 | Means | 21.7 | S.D. | 7.21 |
| | Arith. Concepts | Grade 6 | Means | 16.8 | S.D. | 6.80 |
| | Arith. Applications | Grade 6 | Means | 21.5 | S.D. | 7.83 |
| | Social Studies | Grade 6 | Means | 44.6 | S.D. | 12.41 |
| | Science | Grade 6 | Means | 35.0 | S.D. | 9.69 |
| Subtest: | Paragraph Meaning | Grade 7 | Means | 31.0 | S.D. | 11.86 |
| | Spelling | Grade 7 | Means | 26.4 | S.D. | 10.88 |
| | Language | Grade 7 | Means | 92.5 | S.D. | 17.71 |
| | Arith. Computation | Grade 7 | Means | 19.2 | S.D. | 7.28 |
| | Arith. Concepts | Grade 7 | Means | 16.5 | S.D. | 6.47 |
| | Arith. Applications | Grade 7 | Means | 13.7 | S.D. | 5.14 |
| | Social Studies | Grade 7 | Means | 44.9 | S.D. | 13.15 |
| | Science | Grade 7 | Means | 32.2 | S.D. | 9.82 |
| Subtest: | Paragraph Meaning | Grade 8 | Means | 35.6 | S.D. | 11.95 |
| | Spelling | Grade 8 | Means | 31.0 | S.D. | 12.23 |
| | Language | Grade 8 | Means | 98.0 | S.D. | 17.70 |
| | Arith. Computation | Grade 8 | Means | 23.5 | S.D. | 8.35 |
| | Arith. Concepts | Grade 8 | Means | 20.4 | S.D. | 7.58 |
| | Arith. Applications | Grade 8 | Means | 16.4 | S.D. | 5.76 |
| | Social Studies | Grade 8 | Means | 50.3 | S.D. | 13.52 |
| | Science | Grade 8 | Means | 35.6 | S.D. | 10.09 |
| Subtest: | Paragraph Meaning | Grade 9 | Means | 38.0 | S.D. | 12.25 |
| | Spelling | Grade 9 | Means | 32.9 | S.D. | 12.50 |
| | Language | Grade 9 | Means | 101.5 | S.D. | 18.12 |
| | Arith. Computation | Grade 9 | Means | 25.8 | S.D. | 8.84 |
| | Arith. Concepts | Grade 9 | Means | 22.7 | S.D. | 8.21 |
| | Arith. Applications | Grade 9 | Means | 17.5 | S.D. | 5.68 |
| | Social Studies | Grade 9 | Means | 52.6 | S.D. | 13.60 |
| | Science | Grade 9 | Means | 38.0 | S.D. | 9.98 |

Stanford Achievement Tests

| Primary I Grade I | $r_{11}$ |
|---|---|
| Word Reading | .85 |
| Paragh. Meaning | .90 |
| Vocabulary | .79 |
| Spelling | .92 |
| Word Study Skills | .88 |
| Arithmetic | .95 |

| Primary II Grade 2 | $r_{11}$ |
|---|---|
| Word Meaning | .85 |
| Paragh. Meaning | .93 |
| Sci. Soc. Studies Concepts | .66 |
| Spelling | .94 |
| Word Study Skills | .92 |
| Language | .84 |
| Arith. Comput. | .93 |
| Arith. Concepts | .86 |

(Continued)

Stanford Achievement Tests

| Primary II Grade 3 | $r_{11}$ |
|---|---|
| Word Meaning | .87 |
| Paragraph Meaning | .93 |
| Sci. Soc. Studies Concepts | .79 |
| Spelling | .94 |
| Word Study Skills | .94 |
| Language | .89 |
| Arith. Comput. | .93 |
| Arith. Concepts | .90 |

| Intermediate I Grade 4 | $r_{11}$ |
|---|---|
| Word Meaning | .90 |
| Paragraph Meaning | .92 |
| Spelling | .93 |
| Word Study Skills | .94 |
| Language | .93 |
| Arith. Comput. | .89 |
| Arith. Concepts | .86 |
| Arith. Applications | .86 |
| Social Studies | .86 |
| Science | .91 |

| Intermediate II Grade 5 | $r_{11}$ |
|---|---|
| Word Meaning | .89 |
| Paragraph Meaning | .93 |
| Spelling | .91 |
| Language | .93 |
| Arith. Comput. | .83 |
| Arith. Concepts | .77 |
| Arith. Applications | .84 |
| Social Studies | .89 |
| Science | .88 |

| Intermediate II Grade 6 | $r_{11}$ |
|---|---|
| Word Meaning | .90 |
| Paragraph Meaning | .93 |
| Spelling | .93 |
| Language | .95 |
| Arith. Comput. | .89 |
| Arith. Concepts | .85 |
| Arith. Applications | .89 |
| Social Studies | .91 |
| Science | .90 |

| Advanced Grade 7 | $r_{11}$ |
|---|---|
| Paragraph Meaning | .93 |
| Spelling | .91 |
| Language | .94 |
| Arith. Comput. | .87 |
| Arith. Concepts | .82 |
| Arith. Applications | .77 |
| Social Studies | .89 |
| Science | .88 |

| Advanced Grade 8 | $r_{11}$ |
|---|---|
| Paragraph Meaning | .93 |
| Spelling | .92 |
| Language | .94 |
| Arith. Comput. | .90 |
| Arith. Concepts | .86 |
| Arith. Applications | .81 |
| Social Studies | .91 |
| Science | .89 |

| Advanced Grade 9 | $r_{11}$ |
|---|---|
| Paragraph Meaning | .94 |
| Spelling | .94 |
| Language | .94 |
| Arith. Comput. | .92 |
| Arith. Concepts | .87 |
| Arith. Applications | .79 |
| Social Studies | .90 |
| Science | .90 |

| Grade 9 (Reliability Coefficients) | $r_{11}$ |
|---|---|
| English | .94 |
| Numerical Competence | .90 |
| Math - Part A | .84 |
| Math - Total | .84 |
| Reading | .90 |
| Science - Part A | .87 |
| Science - Total | .87 |
| Social Studies | .85 |
| Spelling | .89 |
| Arts & Humanities | .91 |
| Business & Economics | .88 |
| Tech. Comprehension | .92 |

(Continued)

Stanford Achievement Tests (Reliability Coefficients)

| Grade 10 | $r_a$ | Grade 11 | $r_a$ |
|---|---|---|---|
| English | .94 | English | .94 |
| Numerical Competence | .90 | Numerical Competence | .92 |
| Math - Part A | .89 | Math - Part A | .92 |
| Math - Total | .89 | Math - Total | .93 |
| Reading | .91 | Reading | .92 |
| Science - Part A | .88 | Science - Part A | .87 |
| Science - Total | .89 | Science - Total | .89 |
| Social Studies | .85 | Social Studies | .88 |
| Spelling | .90 | Spelling | .91 |
| Arts & Humanities | .92 | Arts & Humanities | .93 |
| Business & Economics | .88 | Business & Economics | .90 |
| Tech. Comprehension | .93 | Tech. Comprehension | .94 |

| Grade 12 | $r_a$ |
|---|---|
| English | .94 |
| Numerical Competence | .92 |
| Math - Part A | .92 |
| Math - Total | .94 |
| Reading | .91 |
| Science - Part A | .87 |
| Science - Total | .89 |
| Social Studies | .88 |
| Spelling | .92 |
| Arts & Humanities | .93 |
| Business & Economics | .90 |
| Tech. Comprehension | .94 |

| Stanford Achievement Tests | Validity | |
|---|---|---|
| | Content & Construct | Concurrent & Predictive |
| Arith. Applications | Good | Good |
| Arith. Computation | Very Good | Good |
| Arith. Concepts | Very Good | Fair |
| Arith. Total | Good | Fair |
| Language | Very Good | Fair |
| Paragraph Meaning | Very Good | Fair |
| Vocabulary | Very Good | Poor |
| Word Meaning | Very Good | Fair |
| Science | Good | Good |
| Science & Social Studies Concepts | | Good<br>Poor |
| Social Studies | Good | Good |
| Spelling | Very Good | Fair |
| Word Study Skills | Very Good | Fair |

Stanford Achievement Tests (Critiques)

See Also:

Olson, A.V. and Johnson, C.I. "Structure and Predictive Validity of the Frostig Developmental Test of Visual Perception in Grades One and Three." The Journal of Special Education, 1970, 4(1), 49-52.

Stake, R.E. and Hastings, J.T. "The Stanford Achievement Test Battery." Personnel and Guidance Journal, 1964, 43(10), 178-184.

Stanford-Binet Intelligence Scale (1 Form)

Purpose: To test mental ability

Type of Test: Individual intelligence

Range: Two years-Adult

Time for Administration: 60-90 minutes

Cost: $33.20 per set of test materials, $4.40 per 35 record booklets, $4.48 per manual

Subtests: None

Scores: Total IQ, scaled

Nature of Test:

This test differs according to age level. Many of the tests at the lower age levels are quite concrete, dealing with little objects and pictures. At the upper levels, the tests tend to be more abstract and quite heavily verbal. The various tests include tasks calling for display of past learnings, perception of relations, judgment, interpretation, sustained attention, immediate memory, and other cognitive processes.

In testing a child, the examiner begins at a level where the child is likely to succeed, but only with some effort. If the child fails these and appears discouraged, the examiner will drop back to an easier level. Otherwise, he will move ahead, level-by-level, until he reaches a level at which the child fails all tests. When the upper limit has been established, the examiner will go back and establish the level at which the child can do all the tasks.

Stanford-Binet Intelligence Scale (Standardization Sample Scores)

| Age | Means | Standard Deviation |
|---|---|---|
| All | 101.8 | 16.4 |

Reliability: .91
Validity: .82-.90

Stanford-Binet Intelligence Tests (Critiques)

See Also:

Churchhill, W.D. and Smith, S.E. "The Relationship of the 1960 Revised Stanford Binet Intelligence Scale to Intelligence and Achievement Test Scores Over a Three-Year Period." Educational and Psychological Measurement, 1966, 26(4), 1015-1020.

Deloria, D. The Stanford-Binet in the Ypsilanti Curriculum. Why?

Himelstein, P. "Research with the Stanford-Binet, Form L-M: The First Five Years." Psychological Bulletin, 1966, 65(3), 156-164.

Hirshoren, A. "A Comparison of the Predictive Validity of the Revised Stanford-Binet Intelligence Scale and the Illinois Test of Psycholinguistic Abilities," 1969, 35(3), 517-521.

Mueller, M.W. "Validity of Six Tests of Ability with Educable Mental Retardates." Journal of School Psychology, 1968, 6(2), 136-146.

Nelson, C.M. and Hudson, F.G. "Predicting the Reading Achievement of Junior High School EMR Children." American Journal of Mental Deficiency, 1969, 74(3), 415-420.

Scherer, I.W. "The Prediction of Academic Achievement in Brain-Injured Children." Exceptional Children, 1961, 103-107.

Anderson, H.E. and Leton, D.A. "Factor Analysis of the California Test of Mental Maturity." Educational and Psychological Measurements, 1964, 24(3), 513-523.

When the California Test of Mental Maturity (CTMM) was originally published in 1936, it was designed as a group intelligence test, to be the counterpart of the Stanford-Binet individual intelligence test. A short form of the test (CTMM-SF) was published in 1938; both forms were revised in 1963, but their purpose remained the same.

The original factor analysis of the CTMM was carried out in 1938 and the results were utilized in the organization of the subtests on later revisions of the test. The present article is concerned with the factor analytic studies conducted with the 1963 revision of the CTMM and CTMM-SF.

Five levels of the CTMM and CTMM-SF (which collectively span grades one through twelve) were analyzed separately, and means, standard deviations, and Kudar-Richardson reliability coefficients were computed for each of the subtests. The product-moment correlation for each pair of variables was determined and the centroid method was used to factor each of the ten intercorrelation matrices. The factor structures that resulted were rotated by the varimax method and additional necessary rotations and adjustments were made by hand using simple structure and positive manifold as criteria.

The factor analysis revealed three components for the short-form tests: Logical Reasoning, Numerical Reasoning, and Verbal/Memory. The long-form tests also included the Logical and Numerical Reasoning factors. However, three additional factors were determined: Spatial Relations, Memory, and Verbal Concepts. The results for Level 1 (grades one through three) differed somewhat from the others and seemed to indicate the presence of perceptual characteristics in the subtests.

Brim, O.B. "American Attitudes Toward Intelligence Tests." American Psychologists, 20, 1965, 125-130.

This article is an outgrowth of a project under the Russell Sage Foundation which investigated the social consequences of administering standardized ability tests. The author elaborates the various objections which have been raised about testing. He divides the objections into two categories: superficial or apparent criticisms and basic sources of criticisms. The former criticisms are felt to be ones which could be almost alleviated with the proper controls and education. The latter kind of criticisms are harder to alleviate since these deal with more basic issues such as long-range self-esteem, and the issue of equality of all men. He concludes by sayirg: "Our review of these four fundamental sources of opposition to tests leads to the conclusion that psychologists must be concerned about the society in which tests are used. Testing does not occur in isolation; there is always a social context. Test scores have a social meaning. They have impact on man's self-esteem; they influence his life chances; they engage his deepest political and social attitudes. These forces must be understood as part of the social setting in which intelligence testing is carried on."

Churchill, W.D. and Smith, S.E. "The Relationship of the 1960 Revised Stanford-Binet Intelligence Scale to Intelligence and Achievement Test Scores Over a Three-Year Period." Educational and Psychological Measurement, 1966, 26(4), 1015-1020.

The primary objective of this study was to determine the relationship between the 1960 Stanford-Binet Intelligence Scale (S-B) and Lorge-Thorndike IQ's obtained three years later. A secondary objective was to determine the relationship between the 1960 S-B and achievement as measured by the Iowa Test of Basic Skills.

The subjects for the study were 56 members of a seventh grade class in a rural central school. These students had taken the 1960 S-B and the Iowa Test of Basic Skills (Form 2) in third grade, and the Lorge-Thorndike Intelligence Tests and Iowa Tests of Basic Skills (Form 1) in grade 6.

The authors concluded that:

1. The 1960 S-B is a valid measure of "scholastic ability;
2. Group mental tests and comprehensive achievement tests lead to the same prediction;
3. For this small sample, at least, the two IQ tests used could be considered interchangeable.

Deloria, Dennis. "The Stanford-Binet in the Ypsilanti Curriculum. Why?" High Scope Educational Research Foundation.

The author evaluates the Stanford-Binet Intelligence Scale as he defends its use as the key instrument used to assess the affects of the Ypsilanti Cognitively-Oriented Follow Through Curriculum.

The following points are made:

1. The Stanford-Binet (S-B) exhibits high predictive validity for later success in school;

2. The S-B excells in reliability;

3. The S-B measures present ability, not inborn capacity;

4. Though the S-B was normed using middle class white persons, it does predict success in American schools as they currently exist better than any comparable test.

5. The cultural bias of the S-B can be diminished by considering change scores rather than absolute scores;

6. An intelligence test is really only a special kind of achievement test;

7. The variability of S-B scores remains essentially constant as children progress in age.

Gage, G.E. and Naumann, T.F. "Correlation of the Peabody Picture Vocabulary Test and the Wechsler Intelligence Scale for Children." The Journal of Educational Research, 1965, 58(10), 466-468.

The purpose of this study was to determine if the Peabody Picture Vocabulary Test (PPVT) could be used as a reliable and valid measure of intelligence when the subjects were handicapped in sufficient extent or nature to warrant their referral for psychological evaluation.

The data for this study were collected from the records of 30 consecutive referrals to the psychological services department of a public school system. Average age for the total group was 106 months.

Product-moment coefficients of correlation were computed between the IQ's obtained by Form A of the PPVT and the Wechsler Intelligence Scale for Children (WISC) IQ's. Correlations significant at the 1% confidence level were obtained between the PPVT IQ and the Verbal Scale, Performance Scale, and Full Scale IQ's on the WISC.

Thus, the PPVT may be found to be a reliable and valid measure where an assessment of "understanding" vocabulary vs. "speaking" vocabulary is desired.

Himelstein, P. "Research With the Stanford-Binet, Form L-M: The First Five Years." Psychological Bulletin, 1966, 65(3), 156-164.

This article reviews the first five years of research with the 1960 revision of the Stanford-Binet Intelligence Scale (1960-1965). The major areas considered by the reviewer are validity, reliability, sensitivity to extra-test influences, performance of southern Negro children, influence of socio-economic status, diagnostic value, and brief forms.

A major problem in the studies reviewed was the population used; frequently the sample was drawn from retarded or gifted children. This fact decreased the generalizability of the results.

Almost all of the validity studies were concerned with concurrent validity in which the Stanford-Binet was correlated with another test, often a test that just happened to be in the patient's file. The author reports correlations between the Stanford-Binet and the Wechsler Intelligence Scale for Children, the Wechsler Adult Intelligence Scale, the Draw-A-Man Test, the Peabody Picture Vocabulary Test, the Kuhlman Test of Mental Development, the Vineland Social Maturity Scale, the Gesell Developmental Schedule, the California Test of Mental Maturity, the Sequential Test of Educational Progress, the California Achievement Test, and teacher ratings.

The studies revealed little concern for the reliability of the test, the effects of various incentives, and the influence of situational variables, such as "warm" versus "cold" examiners.

In general, the use of the Stanford-Binet short form was found to slightly underestimate the full-scale score. The Stanford-Binet has not been widely employed as a diagnostic tool.

Hirshoren, A. "A Comparison of the Predictive Validity of the Revised Stanford-Binet Intelligence Scale and the Illinois Test of Psycholinguistic Abilities," Exceptional Children, 1969, 35(3), 517-521.

The Illinois Test of Psycholinguistic Ability (ITPA) was developed as a clinical diagnostic test to be used with young children for the identification of specific language disabilities.

The problem of the present study was to determine the predictive validity of the ITPA Total Language Score and of the subtest scores, as well as the predictive validity of the Revised Stanford-Binet IQ, all obtained at the beginning of kindergarten, for school achievement at the beginning of the second grade. For the study, 40 Caucasion children from a heterogeneous neighborhood were tested with the ITPA, the 1960 revision of the Stanford-Binet, and the California Achievement Test (CAT), Lower Primary Form.

The results indicate that the Total Language Score of the ITPA, at least at the kindergarten level, is as valid a predictor as is the Stanford-Binet IQ for school achievement two years later. Performance on the ITPA subtests at the automatic-sequential level is significantly related to performance on all seven of the achievement variables from the CAT. At the representational level of the ITPA, the various subtests are significantly related to various of the CAT subtests. Using stepwise multiple correlation with the ITPA subtests increases the predictive validity of the ITPA. Including the Stanford-Binet IQ as a predictor variable does not improve upon the predictions of future achievement obtained by using the ITPA subtests alone. The ITPA has the added advantage of noting specific areas of language deficits so that remedial programs for individual children may be developed.

Johnson, G.B. "Factors to be Considered in the Interpretation of Intelligence-Test Scores." The Elementary School Journal, 1953, 1-8.

"There are three sources of potential error to be considered in the interpretation of an intelligence quotient. These are to be found in external factors, which relate to the test itself; in internal factors, which include organic and emotional aspects of the individual being tested; and in factors which may be both external and internal and which include the effect of the individual's environment in general." From the detailed analysis of these sources of error, the author concludes:

"1. Select the tests which most nearly suit your purpose in testing. To make the selection requires a knowledge of individual, group, verbal and performance tests and their statistical foundations.

2. Select the tests which have been demonstrated to be both valid and reliable in testing situations similar to your own. This requires a knowledge of the individual and group differences in your situation. Even then, do not put too much credence in the results of one test.

3. If you use standard norms, realize their limitations. If you find them inadequate, prepare local norms.

4. Whichever test you choose, follow closely the directions for administering it. Practice on yourself often before actually administering it.

5. If the test is not machine scored, score it yourself twice.

6. Administer the test under optimum physical and psychological conditions.

7. With both individual and group tests, observe the pupils for factors which may be significant to efficiency of performance, hence to score interpretation.

8. Condition the testees for, not against, intelligence tests. An initial approach in this direction should be to provide pre-service and in-service training for teachers.

9. Understand that intelligence is a composite of factors, in some of which one person may be relatively strong and others relatively weak.

10. Realize that, even with all these precautions, the intelligence test is still in an early phase of development and that the intelligence quotient cannot be considered as precise, as consistent or as true a value as in a unit of measure from the thermometer or the weighing scale."

Kirk, S.A. and McCarthy, J.J. "The Illinois Test of Psycholinguistic Abilities: An Approach to Differential Diagnosis." American Journal of Mental Deficiency, 1961, 66(3), 399-412.

The Illinois Test of Psycholinguistic Abilities (ITPA) was designed by the authors of this article to assess learning dsyfunctions and, in addition, to provide a remedial program for the difficulties thus diagnosed.

The theoretical model given is based upon Osgood's model which is itself an extension of Hull's mediation hypothesis. There are two levels of organization: (1) automatic-sequential (non-meaning) and (2) representational (meaning). There are three psycholinguistic processes:

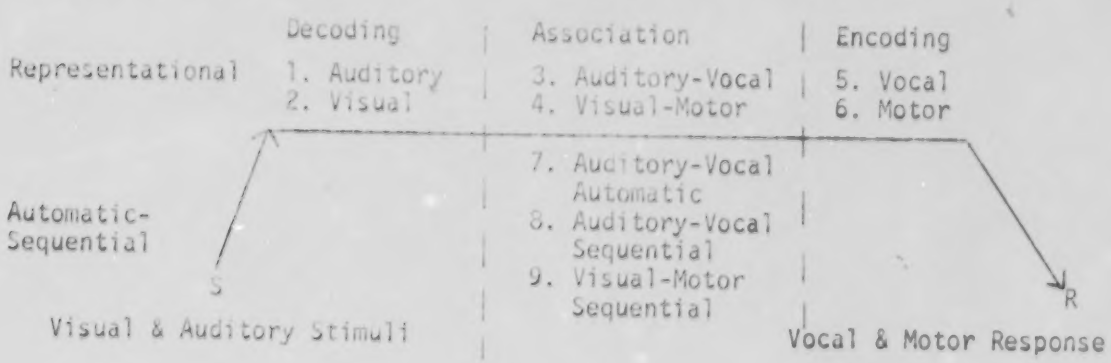
1. Association (ability to manipulate linguistic symbols internally)
2. Encoding
3. Decoding

Four communication channels are utilized:

1. Visual
2. Auditory
3. Motor
4. Vocal

The nine subtests measure the association, encoding and decoding abilities at the representational and automatic-sequential levels, with the tests at the latter level determining disability in less complex tasks. Poor performance on any one of the subtests is construed as the indication of a psycholinguistic defect.

The concept underlying the ITPA is that each clearly defined psycholinguistic area can be tested and those areas that are limiting development can be remediated with specific programs. A clinical model of the nine ITPA subtests is reproduced below:



Knief, L.M. and Stroud, J.B. "Intercorrelations Among Various Intelligence, Achievement and Social Class Scores." Journal of Educational Psychology, 1959, 50(3), 117-120.

In order to obtain information on the relationship of social class to intelligence and achievement test scores, a group of fourth grade students in a Midwestern community were tested over a period of several months. The tests administered were (1) Lorge-Thorndike Intelligence Tests, Verbal; (2) Lorge-Thorndike Intelligence Tests, Nonverbal (Level 3 in both cases); (3) Davis-Eells Games (Elementary Level); (4) Raven's Progressive Matrices; (5) Iowa Test of Basic Skills; and (6) the Warner Index of Status Characteristics.

The following intercorrelations were reported:

| Variable | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | | .309 | .304 | .323 | .179 | .340 |
| 2 | | | .571 | .645 | .405 | .568 |
| 3 | | | | .709 | .437 | .839 |
| 4 | | | | | .521 | .683 |
| 5 | | | | | | .450 |
| 6 | | | | | | |

The authors concluded that their study gave little justification for the use of the Davis-Eells games, Lorge-Thorndike Non-Verbal test, and the Progressive Matrices in conjunction with the Lorge-Thorndike Verbal test for general prediction purposes. However, this information does not negate their usefulness in individual diagnosis.

Pool, D.A. and Brown, R. "The Peabody Picture Vocabulary Test as a Measure of General Adult Intelligence." Journal of Consulting and Clinical Psychology, 1970, 34(1), 8-11.

This study was a search for a rapidly administerable test of general intelligence for adults to be used in large hospitals. The PPVT has not been employed extensively with adults but since it was short, the researchers chose to investigate it. The SS were 150 outpatients referred for psychological evaluation of psychiatric problems, mental retardation, brain damage or potential for vocational rehabilitation. Each S was given the PPVT and also the Dappelt short form of the WAIS. An overall correlation of .81 was found between the two tests. The means were found to be approximately the same in the sub-groups of patients. In the low IQ range (0-59, 60-69), the D WAIS scores were considerably higher than the PPVT. In the higher ranges of IQ the PPVT means were considerably higher than the D-WAIS means; for example, the mean D-WAIS score was 20 points lower than the mean PPVT score. It was shown that although the correlations between the two tests are high, the clinical placement of patients was grossly different for the two tests. Sixty-three percent of the PPVT scores differed by one category or more from the D-WAIS category. Two qualifications were made of these results. (1) The study dealt with PPVT scores only as they predicted D-WAIS scores, and the findings obviously did not bear on the validity of the PPVT; (2) The D-WAIS was not the full WAIS so that there may be differences between the two forms even though they correlate .90. Two major implications were made: (a) comparison of overall means and a visual inspection of a scatter plot of two sets of scores may not readily reveal gross departures from linearity. "This is important to the clinician who deals with individuals and he can't afford to think in terms of overall relations, but must think in terms of 'chances' that a particular score will be grossly misleading;" (b) no matter what is told to parents, teachers and other professionals, there is still a strong tendency for clinicians and psychiatrists to think about IQ scores 'as if' they were WAIS or Stanford Binet scores." In that case, using the PPVT in the lower or upper ranges could do a disservice to the patient.

Rainey, R.G. "A Study of Four School-Ability Tests." The Journal of Experimental Education, 1965, 33(4), 305-319.

The purpose of this study was the evaluation of four school-ability tests: the 1963 versions of the Kuhlman-Anderson, the California Short-Form Test of Mental Maturity, the SRA Primary Mental Abilities Test, and the 1957 form of the Lorge-Thorndike Intelligence Tests. In addition, the Iowa tests of Basic Skills (I.T.B.S.) were administered. The subjects were 2nd, 4th and 7th grade pupils and were designated as belonging to either high, middle or low socio-economic status.

The authors then summarized their results in a series of statements about each test.

The Lorge-Thorndike Test was found to be easy to administer at all levels, from both the students' and testers' viewpoint. The stability of the mean and standard deviation was good at all grade levels in the middle SES area, but poor in the high and low areas. The verbal sections of the test correlated highly with the reading, vocabulary and other language related Iowa Tests; correlations with arithmetic were good, but other tests had better results. The non-verbal section of the test evidenced poor correlations with the Iowa tests when compared with the results obtained from the verbal and total test. The means were highest in the highest SES area and lowest in the low SES area.

The Kuhlman-Anderson Test was easy to administer at all levels and had the added advantage of being short enough to administer during one class period at all grade levels. The mean and standard deviation stability was excellent at all levels and in all SES areas. This test produced the best correlations with the I.T.B.S. in each SES area. The similarity between the means in the high and middle SES areas constituted the major inconsistency of the test.

The California Test of Mental Maturity was not difficult to administer, however, the mean and standard deviation was good only in the high SES area. The language section of the test correlated highly with the verbal subtests of the I.T.B.S. Correlations with arithmetic were higher on the language section than the non-language portions.

The Primary Mental Abilities Test was easy to administer; however, it was the longest of the four tests at each level. The mean and standard deviation stability was the most inconsistent on this test and the greatest variability in all SES areas was produced by this test. The verbal sections of this test correlated well with the verbal subtests of the I.T.B.S. and the total score correlated consistently with all sections of the Iowa Tests. However, none of the Iowa subtests were correlated with the space relations and perceptual speed tests.

As a general conclusion, the authors stated that no one of the tests considered was proven exceptionally inferior or superior.

Stake, R.E. and Hastings, J.T. "The Stanford Achievement Test Battery," Personnel and Guidance Journal, 1964, 43(10), 178-184.

The 1964 edition of the Stanford Achievement Test Battery is stratified in five levels for grades 1-9. At each level, verbal comprehension and skills in rhetoric and quantification are emphasized. For younger pupils there is considerable attention to phonics; for older pupils there are many items devoted to concepts taught in general science, social studies, and advanced arithmetic classes. Determination of whether or not the tests contain appropriate content is in large part a responsibility of local school officials. The Manuals for the Stanford Achievement Tests are well organized to help local officials decide on the appropriateness of content.

The 1964 batteries are longer than the 1963 batteries and the items for the 1964 Edition seem to be more difficult. There is some evidence of the influence of the new curricula, especially in science and mathematics, but school people who have supported innovation in curriculum are likely to have reservations about the item content and the emphasis on grade-equivalent scores.

For the teacher who has to score his own tests, scoring aids packaged with the Stanford Tests are very helpful. The reliability of the sub-tests appears to be on a par with previous edition and they are generally as high as those for any of the currently available, standardized achievement tests. For "knowledge of validity" the authors include a description of the sampling of content, and they put all emphasis upon the user's judgment regarding appropriateness of content for his school.

The arrangement for profiling students' scores on the front of the test booklet or on a separate sheet is attractive and the method of profiling is easy to execute. However, the user should be wary since the use of stanines and the size of intercorrelations may aid over-interpretation.

It seems safe to conclude that if local instructors endorse the content coverage, the Stanford Tests will do as effective a job of measuring elementary school achievement as any standardized battery currently available.

READING ASSESSMENT PROBLEM (READ)

Problem 005 has been deleted from these materials because of intractable problems in simulating data at discrete time periods.

INFORMATION BANK FOR

T QUES (VALIDATION OF A QUESTIONNAIRE)

These materials are meant to be used as abstracts or summaries, highlighting the major points of the articles or presenting only the most relevant normative information. They should be referred to in the same manner as one would refer to a journal such as Psychological Abstracts. More complete information on these sources, or sources not included here, may be found in your library.

This information bank is divided into two sections, as follows: (1) References directly related to this problem; and (2) Descriptive and normative data about variables.

Popham, W. James. The Instructional Objective Exchange: Progress and Prospects. A Symposium Presentation at the American Educational Research Association Annual Meeting. Minneapolis, Minnesota, March 2-6, 1970.

The Instructional Objective Exchange was established in 1968 (September) with the purpose of depositing, developing, and disseminating measurable objectives which might be used by American educators.

This paper is an account of the process of developing the Instructional Objectives Exchange and the instruments for use in measuring teaching effectiveness rather than a description of any of the instruments themselves. Covered are areas such as support monies, distribution of samples, responses, etc., which lead to the establishment of a catalog of available evaluation materials.

It also produced the Instructional Objectives Exchange as a non-profit educational corporation which aims to produce increasingly useful collections of materials which educators can use to rigorously evaluate the quality of their efforts.

Rodin, Miriam and Burton. "Student Evaluations of Teachers," Science, Vol. 177, September 29, 1972, 1164-1166.

The object of this study was to assess the validity of student evaluations by means of a comparison between the objective and subjective criteria of good teaching. The subjective measure consists of asking students to evaluate the teachers' performance. One criticism of this method is that a student's evaluation of an instructor might depend on the grade he receives. There is some evidence that this is true, to some extent. Previous research by Remmers and Elliot into measuring the correlation between objective and subjective criteria yielded inconclusive results, due to the difficulty of defining and controlling the objective criteria (learning).

Method: The instructors were teaching assistants in a large (233 students) undergraduate calculus course. The recitation sections met two days a week. The course content was defined by 40 paradigm problems, which each student had up to 6 opportunities to show mastery of. Grading was objective, and the final grade was based on the total number of problems passed, regardless of how many attempts were made.

Criterion measures: Three measures were obtained for each of the 12 sections: initial ability in calculus (based on mean grade of students in previous quarter), amount learned by the students (mean grade in current quarter), and student evaluation of instructor (mean rating by students).

Results and conclusions: The partial correlation between the objective and subjective measures of teaching ability, with initial ability held constant, was equal to -.746. The instructors with the three lowest subjective scores received the three highest objective scores. The instructor with the highest subjective rating was lowest on the objective measure. There is no obvious explanation for this discrepancy. However, the present data indicate that students are less than perfect judges of teaching effectiveness as defined by objective learning.

Sentag, Maurice, and Pedhazur, Elazar. "Dimensions of Educational Attitudes: Factorial Congruence of Two Scales." Journal of Educational Measurement, 1972, 9(3), 189-198.

Kerlinger's "Attitudes Toward Education Scale" (ES-VII) and Olver and Butcher's "Survey of Opinions About Education" were subjected to first- and second-order factor analyses. After oblique rotations, seven first-order factors emerged. The second order analysis yielded two factors; the first, a bipolar factor involving Kerlinger's conception of traditionalism and Olver and Butcher's radicalism. The results, as far as Kerlinger's scales are concerned, corroborate his earlier findings. Nevertheless, when the operational definition of attitudes toward education is broadened to include Olver and Butcher's scale, the factor structure changes and bipolarity emerges. Consequently, one would have to reject the hypothesis that the different factor structures obtained by Kerlinger and by Olver and Butcher are due to differences in the cultural backgrounds of the respondents. It seems that the two scales represent operational definitions of two theoretical orientations, which, though similar in some respects, are sufficiently different to warrant treating them as distinct orientations. Hence, a choice between the scales will depend on the specific theoretical orientation one wishes to take, and therefore, it is left to the reader to determine which definition of "attitudes toward education" is appropriate for his purposes.

Shoemaker, David M. "Evaluating the Effectiveness of Competing Instructional Programs." Educational Researcher, 1972, 1(5), 5-8.

The difficulty in selecting a particular or single instructional program in education arises because competing programs are constructed frequently on differing sets of objectives and content basis.

Five alternate approaches for assessing the merits of competing instructional programs are described.

(1) Standardized Test.
(2) Specific Test for Each Program.
(3) Common-Unique Objectives:

(a) Objective and content base for each identical across programs.

(4) Common-Unique Objectives.

(b) Only commonality of objectives per se of all programs are considered.

(5) Multiple Matrix Sampling Approach: Stems from recognition that the primary measurement requirement in evaluating competing instructional programs is not individual assessment. An item form is drawn up to cover the range of a particular program. Administration would mean dividing, for example, a 300-item test into a 10 subtests of 30 items each and administering each subtest to a subgroup of examinees selected randomly from those students participating in the program. The results of each test are used to estimate the results that would have to be obtained had all students taken all items in the team population. This procedure makes it possible to estimate accurately the performance of a group of students.

This approach has advantages and disadvantages and in determining its appropriateness involves consideration of setting and desired information.

Warrington, Willard D. The MSU Student Instructional Rating System (SIRS). Paper read at AERA Symposium, April 4, 1972, Chicago.

This is a brief history and report of the current status of Michigan State University's Student Instructional Rating System (SIRS).

SIRS was developed over a two-year period prior to December 1969, when it was adopted by the University Academic Council on a mandatory basis for the total University.

Two doctoral students under the direction of a faculty member undertook to develop an evaluation system and include both faculty and students in the process.

From a series of interviews of faculty and students, preliminary sets of items to be evaluated by students and faculty (randomly selected) were drawn up. These were responded to by students and faculty with 4 yes/no questions. Generally, these were: (1) relevance for evaluation of particular course, (2) would they include the item? (3) need or not for added qualifying information, (4) student's capability to answer question.

From this an experimental form of 56 items was drawn up and administered to a broad sample of classes with 36 different instructors. Five factors were identified and included in a final 28 item form which also included blank spaces for the instructor to insert optional items of his own choosing.

The action of the University Academic counsil required all faculty to use SIRS and make it a tool for the total department's use, i.e., no longer sole property of the individual instructor.

To help in the use of the feedback data and interpretation of scores, SIRS provides an accompanying manual for instructors and consultants.

Graduate Record Examination

Purpose: To assist in graduate school and fellowship candidate selection.

Type of Test: Group aptitude and ability.

Range: College seniors and above.

Time of Administration: 3 hours

Subtests: Verbal and mathematical

Scores: Raw scores reported for each subtest.

Nature of Test:

The G.R.E. Aptitude is similar in format to the familiar College Entrance Examination Board Scholastic Aptitude Test, but pitched at a higher level of difficulty. The verbal score is based upon verbal reasoning and reading comprehension items. The quantitative section includes computation, reasoning, and data interpretation. These items presuppose no more than high school mathematics.

The Aptitude Test-Normative Data by Sex, Based on the Performance of Candidates from May 1967 Through April 1970

| | Percent of Candidates Scoring Lower Than Selected Scaled Scores | | | | | |
|---|---|---|---|---|---|---|
| | Verbal Ability | | | Quantitative Ability | | |
| Score | Men | Women | Total | Men | Women | Total |
| 800 | | | | 98 | | 99 |
| 780 | 99 | 99 | 99 | 97 | | 98 |
| 760 | 98 | 98 | 98 | 94 | 99 | 96 |
| 740 | 97 | 97 | 97 | 91 | 98 | 94 |
| 720 | 95 | 95 | 95 | 88 | 97 | 92 |
| 700 | 93 | 92 | 93 | 83 | 96 | 88 |
| 680 | 90 | 89 | 90 | 80 | 95 | 86 |
| 660 | 87 | 86 | 87 | 75 | 92 | 82 |
| 640 | 84 | 82 | 83 | 70 | 90 | 78 |
| 620 | 80 | 77 | 78 | 65 | 87 | 74 |
| 600 | 75 | 72 | 74 | 60 | 83 | 69 |
| 580 | 70 | 66 | 68 | 54 | 79 | 64 |
| 560 | 65 | 60 | 63 | 49 | 75 | 59 |
| 540 | 58 | 54 | 57 | 43 | 69 | 54 |
| 520 | 53 | 48 | 51 | 38 | 64 | 48 |
| 500 | 47 | 42 | 45 | 34 | 58 | 44 |
| 480 | 41 | 36 | 39 | 29 | 52 | 38 |
| 460 | 35 | 30 | 33 | 24 | 46 | 33 |
| 440 | 29 | 25 | 27 | 20 | 39 | 28 |
| 420 | 24 | 20 | 23 | 16 | 33 | 23 |
| 400 | 19 | 16 | 18 | 13 | 27 | 19 |

(continued)

Percent of Candidates Scoring Lower Than Selected Scaled Scores

| Score | Verbal Ability | | | Quantitative Ability | | |
|---|---|---|---|---|---|---|
| | Men | Women | Total | Men | Women | Total |
| 380 | 15 | 12 | 14 | 10 | 22 | 15 |
| 360 | 12 | 10 | 11 | 7 | 17 | 11 |
| 340 | 9 | 7 | 8 | 5 | 13 | 8 |
| 320 | 7 | 5 | 6 | 4 | 10 | 6 |
| 300 | 5 | 4 | 4 | 2 | 7 | 4 |
| 280 | 3 | 2 | 3 | 1 | 4 | 3 |
| 260 | 2 | 1 | 2 | | 2 | 3 |
| 240 | 1 | | 1 | | 1 | |
| Number of Candidates | 389,474 | 264,408 | 653,882 | 389,443 | 264,388 | 653,831 |
| Mean | 506 | 520 | 512 | 554 | 474 | 522 |
| Stan. Dev. | 125 | 122 | 124 | 132 | 120 | 133 |

Reliability:

Verbal .93
Quantitative .92

Validity: Very Good

INFORMATION BANK FOR

REMEDIAL ARITHMETIC FOR ADULTS

These materials are meant to be used as abstracts or summaries, highlighting the major points of the articles or presenting only the most relevant normative information. They should be referred to in the same manner as one would refer to a journal such as Psychological Abstracts. More complete information on these sources, or sources not included here, may be found in your library.

This information bank is divided into four sections, as follows: (1) References directly related to this problem; (2) Auxiliary References; (3) Descriptive and normative data about variables; and (4) Test critiques, validation studies, etc.

Carman, Robert A. A Systems Approach to the Development of a Junior College Course in Remedial Mathematics, August 1969. (ERIC)

This paper explains the systems approach. Of several definitions, the simplest is that a system is a set of parts coordinated to achieve a (goal or) set of goals. A modified systems approach could likely be an effective tool in educational planning. The information feedback loop permits the manager to base his actions on the actual performance of the system as related to its overall mission and on the performance of any subsystems as related to their objectives. The need constantly arises to determine alternative courses of action, which must be traded off according to system constraints and mission, and to accommodate successive generations of design, directed by re-evaluation as feedback on the initial system becomes available and as constraints and resources change. A course in remedial mathematics was set up under certain constraints: existing programmed material was designed for children; it gave more drill than instruction; other departments needed students with particular math skills; money for change was scarce; new staff could not be expected; no grade, only competence, was required. Its aim was to help students qualify for several departments served by the math department. The author describes the components of this mission by subsystem, and the solution of the problem by the use of a core of courses in arithmetic supplemented by units in algebra, geometry, etc. Other details of the innovation are given. (ERIC)

Losak, John. Do Remedial Programs Really Work? Unpublished paper, 1970.

"This study has demonstrated that for all practical purposes the remedial reading-writing program at Miami-Dade Junior College, as presently designed, does not produce any meaningful differences in student withdrawal from college, is not effective in raising the grade point average during the second semester of college enrollment to a "C" level, does not result in achievement at a higher level in Social Science, Humanities, or English courses, and is not effective in producing a score on a writing test or a reading test that is any higher for those students in the remedial program than it is for those students in a randomly selected Control Group who did not participate in the remedial program.

That there were cases of individuals in the remedial program who improved their scholastic standing and were salvaged academically is attested to by cases of students in the Experimental Group who earned a grade point average of 3.0 ("B") and who improved by as much as 35 percent their standing on a standardized achievement test. Yet, the overall group measures tend to wash out these individual cases when mean scores are considered.

It is precisely the identification of those students who can benefit from a standard remedial program such as offered at Miami-Dade Junior College and elsewhere that is urgently needed. Likewise, those students who cannot benefit from such a standard program must be identified and relevant programs developed for them. It is a gross pedagogical error to assume that the bottom 10 or 15 or 20 percent of a group, based on a standardized achievement test, form a homogeneous group, all capable of benefiting from the same type curriculum.

Modu, Christopher C. "A Description of the Satisfaction Questionnaire for Junior Colleges in Terms of Rotated Factors." Research Bulletin RB-70-17, March 1970, Princeton, New Jersey: ETS, 1970.

"This report describes the dimensions of a Satisfaction Questionnaire administered to junior college students during the experimental phases of the Comparative Guidance and Placement Program of the College Entrance Examination Board. Apart from investigating the stability of their factor structure, the scales were correlated with college grades to determine whether they measure something other than grades. The results produced relatively homogeneous scales indicative of the extent of a student's satisfaction or dissatisfaction with his English courses, Mathematics courses and Major Field and whether he had switched to a lower or more advanced section in English or Mathematics. The scales were also found to bear no relationship with college grades."

Sharon, A.T. "Effectiveness of Remediation in Junior College." Princeton, New Jersey. ETS, Research Bulletin RB-70-50, September 1970.

"Randomly selected groups of students required to take remedial courses at two community colleges were placed in regular courses on an experimental basis. The experimentally placed students were compared with those who enrolled in a regular course after passing a remedial course and with those who did not require remediation. Comparisons were made on ability, interest, motivation, persistence, and performance. The English remedial course had no effect on student satisfaction but produced a small improvement in subsequent performance. The mathematics remedial course eliminated some of the dissatisfaction with the regular course and had a significant effect on subsequent course work. The placement procedures appeared to be more effective in assigning students to appropriate mathematics than to English courses."

Lum, M.K.M. "A Comparison of Under and Overachieving Female College Students," Journal of Educational Psychology, 1960, 51(3), 109-113.

An experimental form of the Brown-Holtzman Survey of Study Habits and Attitudes and a portion of the Dole Vocational Sentence Completion Blank were administered at the University of Hawaii to three groups of Ss (designated as under-normal and overachievers) who were equated for scholastic aptitude and other pertinent variables, but who differed significantly in grades earned. Overachievers have stronger motivation for studying, tend to be more self-confident, and appear to have a greater capacity for working under pressure than underachievers. The latter students show a marked tendency to procrastinate and tend to rely upon external pressures to complete assignments. They are more critical of educational methodology and philosophy than overachievers. Underachievers did not differ from overachievers in their professed study habits. This may suggest to counselors and instructors of classes in "how-to-study" the importance of working on attitudes as well as on the mechanics and conditions of studying.

Norman, R.D. and Daley, M.F. "The Comparative Personality Adjustment of Superior and Inferior Readers." Journal of Educational Psychology, 1959, 50(1), 31-36.

The research literature on reading, reports equivocal results about the relationship between reading ability and personality adjustment. The authors feel that the lack of clear results in the past was due to the inadequacy of statistical methods which could handle multivariate and interactional analysis. Thus, the experiment will take advantage of the newer statistical techniques.

The experiment was conducted in a middle class town. The children used in the study were 6th grade boys with I.Q.'s between 84 and 116 as shown by the California Test of Mental Maturity. Two groups were compared on the California Test of Personality within this group of subjects; a superior reading group which had a mean reading level score of grade 8.1 (N=42) was compared with a poor reading group with a mean grade level reading of 3.9 (N=41). An Anova revealed a significant F ratio on the overall adjustment of the two groups, although the differences between groups on subtests were only quasi-significant. They have hypothesized that several groups of variables seem to be related to poor reading ability although no factor analysis was done. These are poor family interaction, frustration-aggression by others, conflicts over dominance, environmental deprivation, aggressive toward others, impulsivity, rejection of others and inferiority feelings.

Gates Reading Survey

Level of Comprehension Subtest
(Addendum for Remedial Math in the Junior College)

Grade equivalents of scores on the Gates Reading Survey Level of Comprehension Subtest.

| Score | Grade Level |
|---|---|
| 8-11 | 2.9-3.4 |
| 16-19 | 4.3-5.0 |
| 24-27 | 5.8-6.9 |
| 28-31 | 7.2-7.9 |
| 32-35 | 8.2-9.5 |
| 36-39 | 9.9-10.6 |
| 40-43 | 11.1-12.5 |

Gates Reading Tests (3 Forms)

Purpose: To assess reading achievement

Type of Test: Group reading

Range: Grades K-10

Time for Administration: Varies from 40-90 minutes

Cost: Varies from $1.50-$2.75 per 35 tests

Subtests:

Gates Reading Readiness Tests (Grades K-1)

Following Directions
Word Matching
Word-Card Matching
Rhyming
Reading Letters and Numbers

Gates Primary Reading Tests (Grades 1-2.5)

Word Recognition
Paragraph Meaning
Sentence Reading

Gates Advanced Primary Reading Tests (Grades 2.5-3.9)

Word Recognition
Paragraph Meaning

Gates Basic Reading Tests (Grades 3.5-8)

Reading to Appreciate General Significance
Reading to Understand Precise Directions
Reading to Note Details
Reading Vocabulary
Level of Comprehension

Gates Reading Survey (Grades 3.5-10)

Speed and Accuracy
Vocabulary
Level of Comprehension

Scores: Raw; one for each subtest. Gates Reading Survey also includes an accuracy score and a total score.

Nature of Test:

The Gates Reading Tests: The Gates Reading Tests are a unified program for testing reading in the grades. The Gates Reading Readiness Tests are designed to test the degree to which a child is ready to learn to read and also to determine his strength and weakness in each of several important abilities involved in learning to read. These tests are usually given in the early part of the first grade and occasionally near the end of the kindergarten period.

For testing reading in Grade One and the first half of Grade Two, the Gates Primary Reading Tests are provided. These tests measure the level and range of ability in the three most important aspects of reading at the early primary stage. Classes of lower reading ability may use these tests throughout Grade Two.

The Gates Advanced Primary Reading Tests are for use in average classes at midterm and later in Grade Two, and at any time in Grade Three, including the end of the year period. The Advanced Primary series is similar to the Primary series except that it reaches to considerably higher levels of difficulty and will be adequate for able pupils in Grades Two and Three. The Advanced Primary Tests were developed for use at any time during Grade Three. The tests measure range and level of reading vocabulary (Type AWR) and range and level of paragraph comprehension. They do not measure speed of reading.

| Primary Word Recognition | Grade Level | Mean Score | Standard Dev. |
|---|---|---|---|
| | 2.2 | 27.7 | 11.3 |
| | 2.4 | 33.5 | 13.6 |
| | 2.8 | 38.4 | 11.9 |
| Primary Sentence Reading | 2.2 | 23.0 | 11.3 |
| | 2.4 | 26.9 | 13.4 |
| | 2.8 | 37.4 | 11.6 |
| Primary Paragraph Reading | 2.2 | 15.1 | 5.5 |
| | 2.4 | 17.0 | 7.8 |
| | 2.8 | 21.7 | 6.8 |
| Advanced Primary Word Recognition | 3.2 | 24.3 | 9.4 |
| | 3.5 | 29.7 | 9.8 |
| | 3.8 | 32.3 | 9.3 |
| | 4.2 | 36.2 | 9.7 |
| Advanced Primary Paragraph Reading | | | |
| | 3.2 | 14.5 | 5.1 |
| | 3.5 | 19.1 | 4.7 |
| | 3.8 | 20.4 | 4.1 |
| | 4.2 | 24.3 | 6.1 |
| Basic Speed, General Significance | 3.8 | 7.5 | 3.72 |
| | 4.8 | 10.1 | 4.02 |
| | 5.8 | 11.2 | 4.12 |
| | 6.8 | 12.5 | 4.31 |
| | 7.8 | 16.1 | 4.23 |
| | 8.8 | 17.7 | 4.18 |
| Basic Speed, Understanding Directions | 3.8 | 11.1 | 4.82 |
| | 4.8 | 12.2 | 4.32 |
| | 5.8 | 13.6 | 4.41 |
| | 6.8 | 15.1 | 4.85 |
| | 7.8 | 17.2 | 4.34 |
| | 8.8 | 18.0 | 4.36 |
| Basic Speed, Noting Details | 3.8 | 14.1 | 9.91 |
| | 4.8 | 24.5 | 10.10 |
| | 5.8 | 24.3 | 10.70 |
| | 6.8 | 30.5 | 11.61 |
| | 7.8 | 31.5 | 11.32 |
| | 8.8 | 27.5 | 11.21 |

| | Grade Level | Mean Score | Standard Dev. |
|---|---|---|---|
| Basic Vocabulary | 3.8 | 17.4 | 9.2 |
| | 4.8 | 22.8 | 9.4 |
| | 5.8 | 28.5 | 9.6 |
| | 6.8 | 33.6 | 10.2 |
| | 7.8 | 40.6 | 10.1 |
| | 8.8 | 42.7 | 10.3 |
| Basic Level of Comprehension | 3.8 | 16.1 | 7.2 |
| | 4.8 | 22.7 | 8.4 |
| | 5.8 | 29.4 | 8.6 |
| | 6.8 | 34.8 | 8.8 |
| | 7.8 | 43.6 | 8.1 |
| | 8.8 | 46.1 | 7.2 |
| Survey, Speed & Accuracy | 3.8 | 10.3 | 4.6 |
| | 4.8 | 15.5 | 6.1 |
| | 5.8 | 18.4 | 6.4 |
| | 6.8 | 16.1 | 6.4 |
| | 8.8 | 21.8 | 7.1 |
| Survey, Vocabulary | 3.8 | 17.2 | 9.2 |
| | 4.8 | 23.3 | 9.4 |
| | 5.8 | 28.2 | 9.7 |
| | 6.8 | 33.3 | 10.5 |
| | 8.8 | 41.6 | 10.1 |
| Survey, Level of Comprehension | 3.8 | 14.1 | 6.8 |
| | 4.8 | 17.7 | 7.7 |
| | 5.8 | 22.5 | 8.0 |
| | 6.8 | 26.1 | 8.6 |
| | 8.8 | 33.7 | 8.6 |

Gates Reading Tests

Reliability:

| Grade Level | Reliability Coefficient |
|---|---|
| **Primary Word Recognition** | |
| 2.2 | .87 |
| 2.4 | .88 |
| 2.8 | .86 |
| **Primary Sentence Reading** | |
| 2.2 | .86 |
| 2.4 | .87 |
| 2.8 | .87 |
| **Primary Paragraph Reading** | |
| 2.2 | .87 |
| 2.4 | .88 |
| 2.8 | .89 |
| **Advanced Primary Word Recognition** | |
| 3.2 | .84 |
| 3.5 | .88 |
| 3.8 | .85 |
| 4.2 | .86 |

Gates Reading Tests

Reliability:

Advanced Primary Paragraph Reading

| Grade Level | Reliability Coefficient |
|---|---|
| 3.2 | .86 |
| 3.5 | .88 |
| 3.8 | .87 |
| 4.2 | .89 |

Basic Speed, General Significance

| | |
|---|---|
| 3.8 | .82 |
| 4.8 | .84 |
| 5.8 | .83 |
| 6.8 | .84 |
| 7.8 | .82 |
| 8.8 | .83 |

Basic Speed, Understanding Directions

| | |
|---|---|
| 3.8 | .81 |
| 4.8 | .83 |
| 5.8 | .83 |
| 6.8 | .84 |
| 7.8 | .83 |
| 8.8 | .84 |

Basic Speed, Noting Details

| | |
|---|---|
| 3.8 | .84 |
| 4.8 | .86 |
| 5.8 | .86 |
| 6.8 | .87 |
| 7.8 | .86 |
| 8.8 | .85 |

Basic Vocabulary

| | |
|---|---|
| 3.8 | .85 |
| 4.8 | .88 |
| 5.8 | .87 |
| 6.8 | .89 |
| 7.8 | .87 |
| 8.8 | .87 |

Basic Level of Comprehension

| | |
|---|---|
| 3.8 | .86 |
| 4.8 | .87 |
| 5.8 | .88 |
| 6.8 | .89 |
| 7.8 | .86 |
| 8.8 | .87 |

Survey, Speed and Accuracy

| | |
|---|---|
| 3.8 | .82 |
| 4.8 | .84 |
| 5.8 | .87 |
| 6.8 | .83 |
| 8.8 | .86 |

Gates Reading Tests

Reliability:

| Survey, Vocabulary | |
|---|---|
| Grade Level | Reliability Coefficient |
| 3.8 | .84 |
| 4.8 | .89 |
| 5.8 | .86 |
| 6.8 | .88 |
| 8.8 | .87 |

| Survey, Level of Comprehension | |
|---|---|
| 3.8 | .82 |
| 4.8 | .87 |
| 5.8 | .86 |
| 6.8 | .84 |
| 8.8 | .88 |

Validity: Fair to Good

Gates Reading Tests (Critiques)

See Also:

Feldman, S. and Weiner, M. "The Use of a Standardized Reading Achievement Test with Two Levels of Socio-Economic Status Pupils." The Journal of Experimental Education, 1964, 32(3), 269-274.

Fortenberry, W.D. and Broome, B.J. "Comparison of the Gates Reading Survey and the Reading Section of the Wide-Range Achievement Test." Journal of Developmental Reading, 1963, 7(3), 66-68.

## Otis Quick Scoring Mental Ability Test (2 Forms)

Purpose: To measure mental ability--thinking power or the degree of maturity of the mind.

Type of Test: Group intelligence

Range: Grades 1.5-16

Time for Administration: 30-35 minutes

Cost: $2.30-3.20 per 25 tests

Subtests: None

Levels: Alpha Test--Grades 1.5-4
Beta Test--Grades 4-9
Gamma Test--High Schools & Colleges

Scores: Single Overall IQ

Standardized Mean = 100 Standard Deviation = 15

### Nature of Test:

The Alpha Short Form is composed of 45 sets of four pictorial forms representing objects or designs. Each set of four pictures is treated as a nonverbal item, in which the pupil marks the picture which does not belong logically in the group, because it differs in kind.

The Beta and Gamma tests each yield a single score which summarizes the 80 items, including word meaning, verbal analogies, scrambled sentences, interpretation of proverbs, logical reasoning, number series, arithmetic reasoning, and design analogies. More than two thirds of the 80 items of the Beta Test measure some form of verbal competence. There is considerable evidence in the literature on group intelligence tests indicating that a single score reflecting a considerable emphasis on the use of verbal symbols will predict school success as well as a more complex pattern of scores.

School personnel should be especially cautious of applying the test findings to children with learning problems or to those who need remedial help in reading and in the language arts.

### Otis Quick-Scoring Mental Ability Test

| | Reliability | | | Validity | |
|---|---|---|---|---|---|
| | Test Retest | Internal Consistency | Alternate Form | Content & Construct | Concurrent & Predictive |
| Alpha<br>Beta<br>Gamma | .70 | .80-.90 | .70-.80 | Good | Poor |

### Otis Quick-Scoring Mental Ability Tests (Critiques)

See Also:

Lefever, D.W. "Critique of Otis Quick-Scoring Mental Ability Tests." In O.K. Buros (ed.) The Sixth Mental Measurements Yearbook; Highland Park, New Jersey: The Gryphon Press, 1965, 362.

Nelson, C.M. and Hudson, F.G. "Predicting the Reading Achievement of Junior High School EMR Children." American Journal of Mental Deficiency, 1969, 74(3), 415-420.

Yates, A. "Critique of Otis Quick-Scoring Mental Ability Tests." In O.K. Buros (ed.) The Sixth Mental Measurements Yearbook; Highland Park, New Jersey: The Gryphon Press, 1965, 362.

Stanford Achievement Tests

(Addendum for Remedial Math in the Junior College)

Grade equivalents of scores on arithmetic computation test (advanced battery, Gr. 7-9)

| Score | Grade Equivalent | Max. Score Possible |
|---|---|---|
| 2-5 | 2.6-3.9 | 31 |
| 6-9 | 4.2-5.1 | 31 |
| 10-13 | 5.4-6.0 | 31 |
| 14-17 | 6.2-6.8 | 31 |
| 18-21 | 7.2-8.0 | 31 |
| 22-25 | 8.2-8.9 | |
| 26-30 | 9.2-10.8 | |

Maximum possible on concepts subtest = 40
Maximum possible on applications subtest = 36

Stanford-Binet Intelligence Scale (1 Form)

Purpose: To test mental ability

Type of Test: Individual intelligence

Range: Two years-Adult

Time for Administration: 60-90 minutes

Cost: $33.20 per set of test materials, $4.40 per 35 record booklets, $4.48 per manual

Subtests: None

Scores: Total IQ, scaled

Nature of Test:

This test differs according to age level. Many of the tests at the lower age levels are quite concrete, dealing with little objects and pictures. At the upper levels, the tests tend to be more abstract and quite heavily verbal. The various tests include tasks calling for display of past learnings, perception of relations, judgment, interpretation, sustained attention, immediate memory, and other cognitive processes.

In testing a child, the examiner begins at a level where the child is likely to succeed, but only with some effort. If the child fails these and appears discouraged, the examiner will drop back to an easier level. Otherwise, he will move ahead, level-by-level, until he reaches a level at which the child fails all tests. When the upper limit has been established, the examiner will go back and establish the level at which the child can do all the tasks.

Stanford-Binet Intelligence Scale (Standardization Sample Scores)

| Age | Means | Standard Deviation |
|---|---|---|
| All | 101.8 | 16.4 |

Reliability: .91
Validity: .82-.90

Stanford-Binet Intelligence Tests (Critiques)

See Also:

Churchhill, W.D. and Smith, S.E. "The Relationship of the 1960 Revised Stanford Binet Intelligence Scale to Intelligence and Achievement Test Scores Over a Three-Year Period." Educational and Psychological Measurement, 1966, 26(4), 1015-1020.

Deloria, D. The Stanford-Binet in the Ypsilanti Curriculum. Why?

Himelstein, P. "Research with the Stanford-Binet, Form L-M: The First Five Years." Psychological Bulletin, 1966, 65(3), 156-164.

Hirshoren, A. "A Comparison of the Predictive Validity of the Revised Stanford-Binet Intelligence Scale and the Illinois Test of Psycholinguistic Abilities," 1969, 35(3), 517-521.

Mueller, M.W. "Validity of Six Tests of Ability with Educable Mental Retardates." Journal of School Psychology, 1968, 6(2), 136-146.

Nelson, C.H. and Hudson, F.G. "Predicting the Reading Achievement of Junior High School EMR Children." American Journal of Mental Deficiency, 1969, 74(3), 415-420.

Scherer, I.W. "The Prediction of Academic Achievement in Brain-Injured Children." Exceptional Children, 1961, 103-107.

Survey of Study Habits and Attitudes (1 Form)

Purpose: To measure study methods, motivation for studying, and attitudes toward academic activities in the classroom. It is supposed to predict academic achievement in high school and college.

Type of Test: Group

Range: 7th grade through college

Time for Administration: 20-25 minutes

Cost: $3.50 per 25 tests, $1.00 per manual

Subtest: Delay Avoidance
Work Methods
Education Acceptance
Teacher Approval

Scores: One for each subtest. Fifty possible total points for each subtest.

Nature of Test:

The SSHA consists of numbered statements, arranged in two columns per page, to which the student is directed to respond in one of five ways: Rarely (0-15% of the time), Sometimes (16-35%), Frequently (36-65%), Generally (66-85%), or Almost Always (86-100%). Statements for college and high school forms are almost the same, differing only in vocabulary level, age referents, and the like. The basic content and arrangement of both levels are identical. Four 25-item sub-scales are delineated: Work Methods, Delay Avoidance, Teacher Approval, and Education Acceptance. Scores on the first two subscales may be combined to yield a score for Study Habits; similarly, the last two subscales combine to form Study Attitudes. The total score is labeled Study Orientation.

TABLE 1
SSHA-FORM C

PERCENTILE NORMS FOR COLLEGE FRESHMEN
(BASED UPON 3054 CASES FROM NINE DIFFERENT COLLEGES)

| Percentile | Delay Avoid. | Work Methods | Study Habits | Teacher Approv. | Educat. Accept. | Study Attitudes | Study Orient. | Percentile |
|---|---|---|---|---|---|---|---|---|
| 99 | 46-50 | 45-50 | 87-100 | 47-50 | 47-50 | 94-100 | 171-200 | 99 |
| 97 | 44-45 | 42-44 | 83-86 | 46 | 45-46 | 90-93 | 164-170 | 97 |
| 95 | 41-43 | 40-41 | 77-82 | 44-45 | 43-44 | 86-89 | 157-163 | 95 |
| 90 | 38-40 | 37-39 | 72-76 | 42-43 | 41-42 | 82-85 | 149-156 | 90 |
| 85 | 36-37 | 35-36 | 68-71 | 41 | 40 | 79-81 | 144-148 | 85 |
| 80 | 34-35 | 33-34 | 65-67 | 40 | 38-39 | 77-78 | 138-143 | 80 |
| 75 | 32-33 | 32 | 62-64 | 39 | 37 | 75-76 | 134-137 | 75 |
| 70 | 31 | 30-31 | 60-61 | 38 | 36 | 73-74 | 130-133 | 70 |
| 65 | 29-30 | 29 | 57-59 | 36-37 | 35 | 71-72 | 126-129 | 65 |
| 60 | 28 | 28 | 55-56 | 35 | 34 | 69-70 | 122-125 | 60 |
| 55 | 26-27 | 26-27 | 52-54 | 34 | 33 | 67-68 | 118-121 | 55 |
| 50 | 25 | 25 | 50-51 | 33 | 32 | 65-66 | 114-117 | 50 |
| 45 | 23-24 | 24 | 47-49 | 32 | 31 | 63-64 | 110-113 | 45 |
| 40 | 22 | 22-23 | 45-46 | 31 | 30 | 61-62 | 107-109 | 40 |
| 35 | 21 | 21 | 42-44 | 30 | 28-29 | 59-60 | 102-106 | 35 |
| 30 | 19-20 | 20 | 39-41 | 29 | 27 | 56-58 | 98-101 | 30 |
| 25 | 17-18 | 18-19 | 36-38 | 27-28 | 25-26 | 53-55 | 92-97 | 25 |
| 20 | 15-16 | 16-17 | 32-35 | 25-26 | 24 | 50-52 | 85-91 | 20 |
| 15 | 13-14 | 14-15 | 28-31 | 23-24 | 22-23 | 46-49 | 78-84 | 15 |
| 10 | 10-12 | 12-13 | 23-27 | 20-22 | 19-21 | 40-45 | 68-77 | 10 |
| 5 | 7-9 | 9-11 | 18-22 | 17-19 | 15-18 | 34-39 | 59-67 | 5 |
| 3 | 6 | 7-8 | 14-17 | 13-16 | 13-14 | 26-33 | 53-58 | 3 |
| 1 | 0-5 | 0-6 | 0-13 | 0-12 | 0-12 | 0-25 | 0-52 | 1 |
| Mean | 25.0 | 25.1 | 50.1 | 32.7 | 31.4 | 64.1 | 114.2 | Mean |
| SD | 10.0 | 9.2 | 17.5 | 8.0 | 8.3 | 14.8 | 29.7 | SD |

Survey of Study Habits and Attitudes

| Reliability: | | |
|---|---|---|
| | Delay avoidance subtest | .88 |
| | Work methods subtest | .86 |
| | Education acceptance subtest | .83 |
| | Teacher approval subtest | .85 |

## Wide Range Achievement Tests

(1 Form)

Purpose: To assess scholastic achievement

Type of Test: Individual achievement

Range: Kindergarten-college

Time for Administration: 20-45 minutes

Cost: $3.60 per 50 tests

Levels: I-ages 5-11, II-ages 12-24

Subtests: Reading, spelling, arithmetic

Scores: One for each subtest

Nature of Test:

The WRAT was designed to serve as a convenient tool for the study of the basic school subjects and as an adjunct to tests of intelligence and behavior adjustment. The WRAT consists of the following three subtests, at both levels:

1. Reading: recognizing and naming letters and pronouncing words.
2. Spelling: copying marks resembling letters, writing the name, and writing single words to dictation.
3. Arithmetic: counting, reading number symbols, solving oral problems, and performing written computations.

Means & Standard Deviations of Raw Scores of the Reading Test for Ages 5 to Adult--Both Levels

| Age | Mean | Standard Deviation | Age | Mean | Standard Deviation |
|---|---|---|---|---|---|
| 5 | 16.96 | 7.70 | 11 | 38.15 | 11.09 |
| 5 1/2 | 19.37 | 7.75 | 11 1/2 | 40.15 | 11.53 |
| 6 | 24.08 | 10.43 | 12 | 43.92 | 11.78 |
| 6 1/2 | 32.25 | 10.72 | 12 1/2 | 45.64 | 11.81 |
| 7 | 39.85 | 12.30 | 13 | 46.40 | 11.75 |
| 7 1/2 | 46.83 | 12.40 | 13 1/2 | 48.73 | 11.92 |
| 8 | 51.48 | 12.68 | 14 | 51.14 | 11.87 |
| 8 1/2 | 54.17 | 13.04 | 14 1/2 | 52.29 | 12.06 |
| 9 | 57.40 | 13.07 | 15 | 54.31 | 12.13 |
| 9 1/2 | 61.16 | 13.37 | 15 1/2 | 54.93 | 12.20 |
| 10 | 63.18 | 13.44 | 16 | 55.76 | 12.72 |
| 10 1/2 | 66.60 | 13.39 | 17 | 57.29 | 12.76 |
| 11 | 68.48 | 13.35 | 18-19 | 58.74 | 13.79 |
| 11 1/2 | 70.35 | 13.28 | 20-24 | 59.84 | 14.00 |
| 12 | 73.11 | 13.19 | 25-34 | 59.20 | 13.95 |
| 12 1/2 | 77.49 | 12.97 | 35-44 | 59.19 | 14.04 |
| | | | 45 Up | 58.98 | 14.13 |

Means & Standard Deviations of Raw Scores of the Spelling Test for Ages 5 to Adult--Both Levels

| Age | Mean | Standard Deviation | Age | Means | Standard Deviation |
|---|---|---|---|---|---|
| 5 | 13.87 | 5.26 | 11 | 18.40 | 7.58 |
| 5 1/2 | 15.35 | 5.57 | 11 1/2 | 19.58 | 7.71 |
| 6 | 18.70 | 5.70 | 12 | 24.13 | 8.66 |
| 6 1/2 | 23.50 | 5.80 | 12 1/2 | 25.84 | 8.78 |
| 7 | 27.03 | 6.13 | 13 | 27.08 | 8.84 |
| 7 1/2 | 31.15 | 6.30 | 13 1/2 | 29.04 | 8.95 |
| 8 | 33.56 | 6.47 | 14 | 30.35 | 8.73 |
| 8 1/2 | 35.28 | 6.66 | 14 1/2 | 31.02 | 8.82 |
| 9 | 37.56 | 6.91 | 15 | 31.63 | 8.98 |
| 9 1/2 | 40.07 | 7.35 | 15 1/2 | 32.90 | 9.39 |
| 10 | 42.15 | 7.94 | 16 | 33.86 | 9.34 |
| 10 1/2 | 44.05 | 8.01 | 17 | 35.24 | 9.25 |
| 11 | 45.87 | 8.00 | 18-19 | 36.06 | 10.09 |
| 11 1/2 | 46.72 | 8.00 | 20-24 | 36.94 | 10.15 |
| 12 | 48.69 | 7.97 | 25-34 | 35.63 | 10.60 |
| 12 1/2 | 50.81 | 7.76 | 35-44 | 34.86 | 10.89 |
| | | | 45 Up | 33.51 | 10.26 |

Means & Standard Deviations of Raw Scores of the Arithmetic Test for Ages 5 to Adult--Both Levels

| Age | Means | Standard Deviation | Age | Means | Standard Deviation |
|---|---|---|---|---|---|
| 5 | 12.74 | 4.41 | 11 | 19.72 | 4.78 |
| 5 1/2 | 13.51 | 4.49 | 11 1/2 | 21.01 | 4.89 |
| 6 | 17.14 | 4.70 | 12 | 23.71 | 5.46 |
| 6 1/2 | 19.32 | 4.74 | 12 1/2 | 25.22 | 5.70 |
| 7 | 22.16 | 4.62 | 13 | 26.31 | 6.08 |
| 7 1/2 | 24.97 | 4.68 | 13 1/2 | 27.63 | 6.10 |
| 8 | 27.03 | 4.51 | 14 | 28.30 | 6.15 |
| 8 1/2 | 28.50 | 4.42 | 14 1/2 | 29.48 | 6.38 |
| 9 | 29.95 | 4.43 | 15 | 29.50 | 6.37 |
| 9 1/2 | 31.93 | 4.80 | 15 1/2 | 29.65 | 6.63 |
| 10 | 33.72 | 5.77 | 16 | 29.85 | 6.91 |
| 10 1/2 | 37.03 | 6.27 | 17 | 30.60 | 7.25 |
| 11 | 39.48 | 6.50 | 18-19 | 29.12 | 7.70 |
| 11 1/2 | 41.72 | 6.58 | 20-24 | 28.54 | 8.00 |
| 12 | 44.01 | 6.94 | 25-34 | 25.57 | 7.95 |
| 12 1/2 | 46.25 | 6.97 | 35-44 | 23.06 | 7.74 |
| | | | 45 Up | 20.90 | 7.82 |

Reliability Coefficients (r) of the Raw Scores of the Reading Subtest

| Age in Years | Level I r | Age in Years | Level II r |
|---|---|---|---|
| 5 | .981 | 12 | .986 |
| 6 | .986 | 13 | .987 |
| 7 | .993 | 14 | .988 |
| 8 | .991 | 15 | .986 |
| 9 | .989 | 16-17 | .985 |
| 10 | .990 | 18-19 | .983 |
| 11 | .987 | 20-Up | .987 |

Reliability of the Spelling Subtest

| Age in Years | Level I r | Age in Years | Level II r |
|---|---|---|---|
| 5 | .971 | 12 | .979 |
| 6 | .963 | 13 | .989 |
| 7 | .977 | 14 | .987 |
| 8 | .978 | 15 | .982 |
| 9 | .977 | 16-17 | .984 |
| 10 | .981 | 18-19 | .984 |
| 11 | .982 | 20-Up | .986 |

Reliability Coefficients (r) of the Raw Scores of the Arithmetic Subtest

| Age in Years | Level I r | Age in Years | Level II r |
|---|---|---|---|
| 5 | .966 | 12 | .940 |
| 6 | .959 | 13 | .957 |
| 7 | .962 | 14 | .956 |
| 8 | .948 | 15 | .966 |
| 9 | .942 | 16-17 | .955 |
| 10 | .948 | 18-19 | .969 |
| 11 | .945 | 20-Up | .970 |

Wide Range Achievement Tests

| Validity: | Content & Construct | Concurrent & Predictive |
|---|---|---|
| Arithmetic | *Good | Good |
| Reading | Good | Good |
| Spelling | Good | Good |

*Kindergarten - Very Good

Wide Range Achievement (Critiques)

See Also:

Fortenberry, W.D. and Broome, B.J. "Comparison of the Gates Reading Survey and the Reading Section of the Wide-Range Achievement Test." Journal of Developmental Reading, 1963, 7(3), 66-68.

Hopkins, K.D., Dobson, J.C. and Oldridge, O.A. "The Concurrent and Congruent Validities of the Wide-Range Achievement Test." Educational and Psychological Measurement, 1962, 22(4), 791-793.

Mueller, M.W. "Validity of Six Tests of Ability with Educable Mental Retardates." Journal of School Psychology, 1968, 6(2), 136-146.

Brim, O.B. "American Attitudes Toward Intelligence Tests." American Psychologists, 20, 1965, 125-130.

This article is an outgrowth of a project under the Russell Sage Foundation which investigated the social consequences of administering standardized ability tests. The author elaborates the various objections which have been raised about testing. He divides the objections into two categories: superficial or apparent criticisms and basic sources of criticisms. The former criticisms are felt to be ones which could be almost alleviated with the proper controls and education. The latter kind of criticisms are harder to alleviate since these deal with more basic issues such as long-range self-esteem, and the issue of equality of all men. He concludes by saying: "Our review of these four fundamental sources of opposition to tests leads to the conclusion that psychologists must be concerned about the society in which tests are used. Testing does not occur in isolation; there is always a social context. Test scores have a social meaning. They have impact on man's self-esteem; they influence his life chances; they engage his deepest political and social attitudes. These forces must be understood as part of the social setting in which intelligence testing is carried on."

Churchill, W.D. and Smith, S.E. "The Relationship of the 1960 Revised Stanford-Binet Intelligence Scale to Intelligence and Achievement Test Scores Over a Three-Year Period." Educational and Psychological Measurement, 1966, 26(4), 1015-1020.

The primary objective of this study was to determine the relationship between the 1960 Stanford-Binet Intelligence Scale (S-B) and Lorge-Thorndike IQ's obtained three years later. A secondary objective was to determine the relationship between the 1960 S-B and achievement as measured by the Iowa Test of Basic Skills.

The subjects for the study were 56 members of a seventh grade class in a rural central school. These students had taken the 1960 S-B and the Iowa Test of Basic Skills (Form 2) in third grade, and the Lorge-Thorndike Intelligence Tests and Iowa Tests of Basic Skills (Form 1) in grade 6.

The authors concluded that:

1. The 1960 S-B is a valid measure of "scholastic ability;
2. Group mental tests and comprehensive achievement tests lead to the same prediction;
3. For this small sample, at least, the two IQ tests used could be considered interchangeable.

Deloria, Dennis. "The Stanford-Binet in the Ypsilanti Curriculum. Why?" High Scope Educational Research Foundation.

The author evaluates the Stanford-Binet Intelligence Scale as he defends its use as the key instrument used to assess the affects of the Ypsilanti Cognitively-Oriented Follow Through Curriculum.

The following points are made:

1. The Stanford-Binet (S-B) exhibits high predictive validity for later success in school;

2. The S-B excells in reliability;

3. The S-B measures present ability, not inborn capacity;

4. Though the S-B was normed using middle class white persons, it does predict success in American schools as they currently exist better than any comparable test.

5. The cultural bias of the S-B can be diminished by considering change scores rather than absolute scores;

6. An intelligence test is really only a special kind of achievement test;

7. The variability of S-B scores remains essentially constant as children progress in age.

Eagle, N. and Harris, A.S. "Interaction of Race and Test on Reading Performance Scores." Journal of Educational Measurement, 1966, 6(3), 131-135.

This study examines the relationship between race and performance on two nationally standardized reading tests. The appropriate reading tests of the Iowa Tests of Basic Skills and Metropolitan Achievement Battery were administered to all fourth and sixth grade students in all elementary schools of an urban school district near New York City. Although white pupils earned higher scores than non-white pupils on both tests, the Metropolitan produced significantly greater differences between the races than the Iowa, at both grade levels. White upper elementary children were "favored" by the Metropolitan and Negro children were "favored" by the Iowa.

The authors feel that this finding is of importance since, if an interaction exists between social class or ethnic group and reading test, knowledgeable administrators could use combinations of tests to demonstrate either (a) maximum school reading performance for public relations purposes, or (b) minimal school reading performance, possibly to qualify for more Federal or state compensatory program aid.

Feldman, S. and Weiner, M. "The Use of a Standardized Reading Achievement Test with Two Levels of Socio-Economic Status Pupils." The Journal of Experimental Education, 1964, 32(3), 269-274.

The authors are concerned that, while most tests are standardized for normally distributed populations, the populations most often actually tested are homogeneous. Lower and middle socio-economic level fifth grade students were tested with the Gates Advanced Primary Reading Test and the Gates Basic Reading Test, Level of Comprehension subtest.

The purpose of this study was an analysis of these two tests in an effort to determine comparability and the effect of socio-economic level on test scores. After meeting several difficulties in attempting to use such heterogeneously standardized tests for evaluating the homogeneous groups of subjects, the authors made the following recommendations concerning the use of standardized tests in assessing the reading skills of students in different socio-economic levels:

1. Even though test publishers state that grade scores are comparable in adjacent tests in a battery of reading achievement tests, two points must be considered: (a) the highest and lowest possible grade scores for each test; and (b) whether both tests may be measuring different reading skills. When two different ability groups are given the test, then factors a and b above become even more crucial.

2. Tests should be tried out with a small pilot-group before any large-scale test administration is undertaken.

3. If the purpose of the test user is to distinguish among the pupils' skills measured by the test, then the use of a standardized test with a homogeneous group will very likely not provide the desired results.

Fortenberry, W.D. and Broome, B.J. "Comparison of the Gates Reading Survey and the Reading Section of the Wide-Range Achievement Test." Journal of Developmental Reading, 1963, 7(3), 66-68.

This study looked at the relationship between the group test, the Gates Reading Survey and the individual test, the Wide-Range Achievement Test. The researchers were interested in discovering whether the group test could predict the individual test, thus allowing them to use the group test as a basis for diagnosis rather than having to administer the Wide-Range Achievement Test. High correlations were obtained between the two tests when administered to 98 fourth grade pupils enrolled in elementary schools in Lousiana. The students had IQ's ranging from 83 to 126 with a mean of 103. They came from lower to upper class families. The correlations between GRS section on Speed of reading and the WRA Reading Section was .68; the correlation between the spelling section of the GRS and the WRA reading section was .80; and the correlation between the GRS comprehension section and the WRA reading section was .67. It was concluded that the GRS could be used as a basis for predicting the WRA.

Higgins, J.H. "Critique of the Survey of Study Habits and Attitudes," Journal of Counseling Psychology, 1967(14), 392-393.

The purpose of the SSHA is to identify students whose study habits and attitudes are different from the high-achieving student. It is also designed to aid in helping students with academic difficulties by providing a basis for the help needed.

The author says that there is considerable correlation among the subtests so that the unique variance of each subtest is considerably reduced. The reliability and validity are adequate. It is best used by counselors in an aid in high schools, junior colleges and four-year colleges.

Himelstein, P. "Research With the Stanford-Binet, Form L-M: The First Five Years." Psychological Bulletin, 1966, 65(3), 156-164.

This article reviews the first five years of research with the 1960 revision of the Stanford-Binet Intelligence Scale (1960-1965). The major areas considered by the reviewer are validity, reliability, sensitivity to extra-test influences, performance of southern Negro children, influence of socio-economic status, diagnostic value, and brief forms.

A major problem in the studies reviewed was the population used; frequently the sample was drawn from retarded or gifted children. This fact decreased the generalizability of the results.

Almost all of the validity studies were concerned with concurrent validity in which the Stanford-Binet was correlated with another test, often a test that just happened to be in the patient's file. The author reports correlations between the Stanford-Binet and the Wechsler Intelligence Scale for Children, the Wechsler Adult Intelligence Scale, the Draw-A-Man Test, the Peabody Picture Vocabulary Test, the Kuhlman Test of Mental Development, the Vineland Social Maturity Scale, the Gesell Developmental Schedule, the California Test of Mental Maturity, the Sequential Test of Educational Progress, the California Achievement Test, and teacher ratings.

The studies revealed little concern for the reliability of the test, the effects of various incentives, and the influence of situational variables, such as "warm" versus "cold" examiners.

In general, the use of the Stanford-Binet short form was found to slightly underestimate the full-scale score. The Stanford-Binet has not been widely employed as a diagnostic tool.

Hirshoren, A. "A Comparison of the Predictive Validity of the Revised Stanford-Binet Intelligence Scale and the Illinois Test of Psycholinguistic Abilities," Exceptional Children, 1969, 35(3), 517-521.

The Illinois Test of Psycholinguistic Ability (ITPA) was developed as a clinical diagnostic test to be used with young children for the identification of specific language disabilities.

The problem of the present study was to determine the predictive validity of the ITPA Total Language Score and of the subtest scores, as well as the predictive validity of the Revised Stanford-Binet IQ, all obtained at the beginning of kindergarten, for school achievement at the beginning of the second grade. For the study, 40 Caucasion children from a heterogeneous neighborhood were tested with the ITPA, the 1960 revision of the Stanford-Binet, and the California Achievement Test (CAT), Lower Primary Form.

The results indicate that the Total Language Score of the ITPA, at least at the kindergarten level, is as valid a predictor as is the Stanford-Binet IQ for school achievement two years later. Performance on the ITPA subtests at the automatic-sequential level is significantly related to performance on all seven of the achievement variables from the CAT. At the representational level of the ITPA, the various subtests are significantly related to various of the CAT subtests. Using stepwise multiple correlation with the ITPA subtests increases the predictive validity of the ITPA. Including the Stanford-Binet IQ as a predictor variable does not improve upon the predictions of future achievement obtained by using the ITPA subtests alone. The ITPA has the added advantage of noting specific areas of language deficits so that remedial programs for individual children may be developed.

Hopkins, K.D., Dobson, J.C. and Oldridge, O.A. "The Concurrent and Congruent Validities of the Wide-Range Achievement Test." Educational and Psychological Measurement, 1962, 22(4), 791-793.

The WRAT was first published in 1946. The reviews of the test were unfavorable because of its questionable validity and norming procedures, and because of its unsubstantiated claims of clinical symptoms suggested by the authors. Testing authorities have not examined the test since that time although it is the second most used achievement test in Psychological clinics. The reason for this seems to lie in the fact that it is easy to administer and it eliminates almost all of the chance variance.

This study compared the WRAT scores on reading achievement with teacher ratings of children's reading ability and the reading sections on the California Achievement Test. This would show the concurrent validity of the WRAT. It was found that with a sample of 501 pupils in grades 1-5, the WRAT and CAT showed good concurrent validity with correlations of from .851 to .906 with teacher ratings and concurrent validity with the CAT of .739 to .856. The results tend to validate at least one portion of the WRAT.

Johnson, G.B. "Factors to be Considered in the Interpretation of Intelligence-Test Scores." The Elementary School Journal, 1953, 1-8.

"There are three sources of potential error to be considered in the interpretation of an intelligence quotient. These are to be found in external factors, which relate to the test itself; in internal factors, which include organic and emotional aspects of the individual being tested; and in factors which may be both external and internal and which include the effect of the individual's environment in general." From the detailed analysis of these sources of error, the author concludes:

"1. Select the tests which most nearly suit your purpose in testing. To make the selection requires a knowledge of individual, group, verbal and performance tests and their statistical foundations.

2. Select the tests which have been demonstrated to be both valid and reliable in testing situations similar to your own. This requires a knowledge of the individual and group differences in your situation. Even then, do not put too much credence in the results of one test.

3. If you use standard norms, realize their limitations. If you find them inadequate, prepare local norms.

4. Whichever test you choose, follow closely the directions for administering it. Practice on yourself often before actually administering it.

5. If the test is not machine scored, score it yourself twice.

6. Administer the test under optimum physical and psychological conditions.

7. With both individual and group tests, observe the pupils for factors which may be significant to efficiency of performance, hence to score interpretation.

8. Condition the testees for, not against, intelligence tests. An initial approach in this direction should be to provide pre-service and in-service training for teachers.

9. Understand that intelligence is a composite of factors, in some of which one person may be relatively strong and others relatively weak.

10. Realize that, even with all these precautions, the intelligence test is still in an early phase of development and that the intelligence quotient cannot be considered as precise, as consistent or as true a value as in a unit of measure from the thermometer or the weighing scale."

Knief, L.M. and Stroud, J.B. "Intercorrelations Among Various Intelligence, Achievement and Social Class Scores." Journal of Educational Psychology, 1959, 50(3), 117-120.

In order to obtain information on the relationship of social class to intelligence and achievement test scores, a group of fourth grade students in a Midwestern community were tested over a period of several months. The tests administered were (1) Lorge-Thorndike Intelligence Tests, Verbal; (2) Lorge-Thorndike Intelligence Tests, Nonverbal (Level 3 in both cases); (3) Davis-Eells Games (Elementary Level); (4) Raven's Progressive Matrices; (5) Iowa Test of Basic Skills; and (6) the Warner Index of Status Characteristics.

The following intercorrelations were reported:

| Variable | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | | .309 | .304 | .323 | .179 | .340 |
| 2 | | | .571 | .645 | .405 | .568 |
| 3 | | | | .709 | .437 | .839 |
| 4 | | | | | .521 | .683 |
| 5 | | | | | | .450 |
| 6 | | | | | | |

The authors concluded that their study gave little justification for the use of the Davis-Eells games, Lorge-Thorndike Non-Verbal test, and the Progressive Matrices in conjunction with the Lorge-Thorndike Verbal test for general prediction purposes. However, this information does not negate their usefulness in individual diagnosis.

Lefever, D.W. "Critique of Otis Quick-Scoring Mental Ability Test," in O.K. Buros (ed.) The Sixth Mental Measurements Yearbook; Highland Park, New Jersey; The Gryphon Press, 1965, 362.

The reviewer has several criticisms of the new edition of the Otis Quick-Scoring Mental Ability Tests, many of which are concerned with the rather vague and incomplete manner in which the norms were derived. Other than this, he feels that the most serious weaknesses of the revised tests appear to be the lack of percentile norm tables of any description and the failure to furnish normative data on the comparability of the two new forms for each level. The manuals are very brief and the reviewer feels that detailed suggestions for teachers, counselors and supervisors on the application of the tests to instructional and guidance problems would be helpful. More recent evidence concerning the validity of Gamma is to be desired, as well as more extensive recent data for the predictive power of both Beta and Gamma tests.

The reviewer does note, however, that most of his criticisms are directed at the amount of information furnished to the user of the tests rather than against the quality or value of the tests themselves. The Otis Quick-Scoring Mental Ability Tests, as the title implies, do furnish a short and easily scored indicator of scholastic aptitude. Such a measure, if interpreted with care, can be useful to both teacher and counselor by revealing within fairly broad limits of accuracy the probable level of academic achievement for a majority of pupils.

Popham, W.J. and Moore, M.R. "A Validity Check on the Brown-Holtzman Survey of Study Habits and Attitudes and the Borow College Inventory of Academic Adjustments, Personnel and Guidance Journal, 1960, 38, 552-554.

Students at Kansas State College in Pittsburg, Kansas, were given the SSHA and the CIAA before they entered a general psychology class. Correlations were obtained between their grade point average and their test scores.

The present study was designed to test the hypothesis that college students' scores on the SSHA and CIAA evidence a significantly stronger relationship to students' grades than they do to an index of students' intellectual aptitude. The SSHA and CIAA were administered to 287 undergraduate students. Data regarding subjects' ACE scores and grade ratios earned during the previous semester were obtained. Obtained correlation coefficients and assessment of differences between these coefficients revealed a significantly stronger relationship between subjects' scores on either instrument and grades than between students' scores on either instrument and ACE scores. It was concluded that the SSHA and CIAA do, in fact, measure something other than intellectual aptitude which is associated with student academic success and may, therefore, aid the counselor in identifying potential academic overachievers and underachievers.

Roark, A.E. and Harrington, S.A. "Review of the Survey of Study Habits and Attitudes," Journal of Educational Measurement, 6(2), pp. 120-122, 1969.

The authors of the SSHA claim it has a myriad of uses. The authors of this critique feel that the SSHA has limited use because of some inherent weaknesses.

The scores of the SSHA are easily manipulated by the student, thus it is susceptible to faked scores. The self-respect of the student is the only criterion and this is subject to great variation. The manual claims that this test could be used as teaching aid but does not give any details on how this is to be done. There is only a moderate relationship between grade point average and the SSHA, and there are no studies which show its importance in academic success.

It is a well-constructed test and could be used well by counselors.

Shay, C.B. "Critique of the S.S.H.A." In O.K. Buros (ed.) The Seventh Mental Measurements Yearbook, Vol. II, 1970, 782-783.

The SSHA has been revised from the 1953 edition. The test is now better than it was.

"Evidence for validity is presented in terms of low correlation coefficients (mean .21 for C, .27 for H) between SSHA and aptitude tests and higher correlations (mean .36 for C, .49 for H) between SSHA and grades, which suggest that the test is related to grades but is not just a measure of ability. Partial correlation coefficients (average .43) between SSHA Form H and grades, with aptitude level held constant, also support this notion. In addition, multiple correlation coefficients combining SSHA with aptitude test scores show that the correlation with grades is increased from .07 to .16 over the correlation of grades with aptitude scores alone. Although these are moderate increases, they are in a promising direction; however, no evidence is given for crossvalidation of these results.

Intercorrelation coefficients of SSHA sub-scale scores range from .51 to .75, minimally supportive of the success of the attempt to create unique sub-scales of use to counselors. Counselors requested these sub-scales; it is hoped these high correlations will deter them from considering the sub-scales measures of unique traits.

As pointed out by previous reviewers, the most troublesome aspect of SSHA validity deals with the inventory's use as a predictor. The basic assumption underlying the inventory is that some students earn poor grades because of poor study habits and attitudes, and that if these habits and attitudes can be identified and changed, their grades will improve. The prediction which follows is that students with poor habits and attitudes will earn low grades, and those with good habits and attitudes will earn high grades; thus, the inventory can select those students who will earn good grades. The trouble with this assumption is that a student may know proper study habits and attitudes but fail to employ them, and besides, there are many other reasons for good and poor grades. Validation studies cited in the manual did not use the SSHA as a selection instrument, and in view of its complete dependence upon frankness of responding, the inventory might never be used in this way. Studies of its usefulness in counseling, particularly those which show that a change of study habits and attitudes results in a change of grades, have yet to be conducted. Until then, its ultimate justification for use must rest on unverified assumptions.

The SSHA is easy to use and to score; directions are clear and complete. Each answer sheet includes a diagnostic profile, which can be filled in by the student. Instructions for completing this profile call for blackening subtest percentile bars as deviations from the median. A student can readily gain the impression that scores below the median are "low" and scores above are "high." No discussion is given about how large differences need to be, either between subscales or in comparison with the norm group, in order to be important.

In summary, the Survey of Study Habits and Attitudes has been carefully devised and has satisfactory reliability; there is statistical evidence for some of the least important aspects of validity. Its use as a predictor is unwarranted, and subscale scores should be interpreted with caution. Additional research is needed to justify its use as a screening or diagnostic instrument. Nonetheless, it is a good teaching aid for teachers and counselors--and useful to those students who are frank in responding and motivated to improve.

Stake, R.E. and Hastings, J.T. "The Stanford Achievement Test Battery," Personnel and Guidance Journal, 1964, 43(10), 178-184.

The 1964 edition of the Stanford Achievement Test Battery is stratified in five levels for grades 1-9. At each level, verbal comprehension and skills in rhetoric and quantification are emphasized. For younger pupils there is considerable attention to phonics; for older pupils there are many items devoted to concepts taught in general science, social studies, and advanced arithmetic classes. Determination of whether or not the tests contain appropriate content is in large part a responsibility of local school officials. The Manuals for the Stanford Achievement Tests are well organized to help local officials decide on the appropriateness of content.

The 1964 batteries are longer than the 1963 batteries and the items for the 1964 Edition seem to be more difficult. There is some evidence of the influence of the new curricula, especially in science and mathematics, but school people who have supported innovation in curriculum are likely to have reservations about the item content and the emphasis on grade-equivalent scores.

For the teacher who has to score his own tests, scoring aids packaged with the Stanford Tests are very helpful. The reliability of the sub-tests appears to be on a par with previous edition and they are generally as high as those for any of the currently available, standardized achievement tests. For "knowledge of validity" the authors include a description of the sampling of content, and they put all emphasis upon the user's judgment regarding appropriateness of content for his school.

The arrangement for profiling students' scores on the front of the test booklet or on a separate sheet is attractive and the method of profiling is easy to execute. However, the user should be wary since the use of stanines and the size of intercorrelations may aid over-interpretation.

It seems safe to conclude that if local instructors endorse the content coverage, the Stanford Tests will do as effective a job of measuring elementary school achievement as any standardized battery currently available.

Yates, A. "Critique of Otis Quick-Scoring Mental Abilities Tests," in O.K. Buros (ed.) The Sixth Mental Measurements Yearbook; Highland Park, New Jersey: The Gryphon Press, 1965, 362.

This new form comprises three tests: the Alpha Short Form is designed for grades 1-4; Beta, suitable for grades 4-9, is a revision of the Intermediate Examination of the Otis Self-Administering Tests of Mental Ability; and Gamma, intended for high schools and colleges, is a revision of the Higher Examination of the older series.

The avowed purpose of the tests is to measure mental ability defined as 'thinking power or the degree of maturity of the mind.' In making up the tests the author's aim was to choose items which depended as little as possible on schooling, except that in the Beta and Gamma tests questions on vocabulary and arithmetic reasoning, recognized as being largely measures of achievement, are included.

The reliability coefficients quoted for these tests are not strikingly high. Therefore, the results of applying such a test should be interpreted cautiously. The evidence concerning the validity of the test for predictive purposes is somewhat slender. It is doubtful whether, for the normal purposes of classification and guidance within a school, teachers are likely to find the tests more serviceable than measures of attainment and their own judgments of their pupils' progress. The tests are obviously serviceable instruments, however, for the purposes of course classification when this had to be carried out without access to relevant information about children's previous educational progress.

# INFORMATION BANK FOR
# BUSING

These materials are meant to be used as abstracts or summaries, highlighting the major points of the articles or presenting only the most relevant normative information. They should be referred to in the same manner as one would refer to a journal such as Psychological Abstracts. More complete information on these sources, or sources not included here, may be found in your library.

This information bank is divided into four sections, as follows: (1) References directly related to this problem; (2) Auxiliary References; (3) Descriptive and normative data about variables; and (4) Test critiques, validation studies, etc.

Barr, Jo. "A Resume of the Findings of the Follow-Through Project Evaluation 1969-70." (ERIC)

This report summarizes the findings of the 1969-70 evaluation of Project Follow-Through in Wichita, Kansas. In the first year of the program, 186 Headstart graduates were bused from low income residential areas to four elementary schools located on the periphery of the city. The receiving schools varied in socio-economic status of pupils from middle class to working class. The 1969-70 school year completed the second year of Wichita's Follow-Through Project. Students were evaluated by questionnaires given to teachers: in addition, to measure pupil adjustment and achievement, tests such as the Cognitive Abilities Test, the California Test of Personality, and the Illinois Test of Psycholinguistic Abilities were given. Parent participation was encouraged. The results of the Cognitive Abilities Test indicate that though Follow-Through students scored higher than Headstart students at pretest, by post-test there were no significant differences; both groups, however, made significant gains. In the Illinois Test of Psycholinguistic Abilities, the experimental Follow-Through students, after two years, scored significantly higher than control groups. It is held that the program had some good results. (ERIC)

BENEFITS OF K-3, 4-6 STRUCTURE FOR ELEMENTARY SCHOOLS

Prepared by Feasibility Committee

The K-3, 4-6 plan was not adopted by the Board of Education simply because it provided the most efficient method of desegregating the elementary schools, but also because it provided educational advantages not offered by other plans. For example:

--By reducing the number of grades in each school, the number of children at each grade level is increased. This means that specialized personnel can be put to better use, as can specialized equipment.

--The formation of 4-6 schools allows for easier transition to junior high school. Because there will be fewer schools feeding into the existing junior high schools, this will allow for closer staff cooperation between the 4-6 schools and the staff of the junior high schools.

--Class size will be more nearly equalized throughout the city due to the presence of more students of the same grade level in each school.

--Specialized facilities such as library book collections can be geared to specific grade level.

--A greater variety will be allowed in the grouping and scheduling of children. With more teachers of the same grade level in a given school, more stress can be placed on the needs of individual pupils.

--The formation of each of the zones will provide a natural common interest among adults in the zone and can be a spur to communication among parents of different backgrounds.

--The movement of teachers to match the new grade groupings will allow for greater racial integration of the respective school staffs. It is important, for educational reasons, that all our youngsters be taught by teachers with a wide variety of racial backgrounds.

Bograd, Judith. "Busing: Who Didn't Want Their Children Along for the Ride?" NCRIEEO Newsletter, Vol. 3, No. 2, May 1972, 14-15.

Objections to Transportation (1930)

Some of the objections that are usually raised against transportation of pupils are:

1. The time spent on the road going to and coming from school.
2. The danger of physical injury to the pupils.
3. Social relationship of pupils on buses.
4. Cost of transportation.
5. Pupils will be late getting to road and miss the bus.
6. Children have no place to wait for bus after they get to the road, if the weather is bad.
7. Pupils may get sick away from home.
8. Buses unable to get over the roads in bad weather.
9. Attendance will be poor at school on account of transportation.

Although most of the objections are imaginary rather than real, some of them are vital. They may be answered in many ways; some of the most important answers are:

1. Of course some poeple must live on the end of the route and their children will be later than the others in getting home. Sometimes this is an hour or more after school closes. There is no alternative if the children are to be given standard school advantages. In most instances the children will reach home sooner from the consolidated school than they would by walking home from the local small school.
2. With the increased number and the increased speed of cars and trucks on our highways today it is much more dangerous for children to walk down the highway than to ride in a bus. If proper care is taken in the selection of a bus that is built with the safety of the pupils in mind, there is practically no danger to pupils riding in it.
3. Where a great number of children or adults are put in the close relationship that they are in a bus there is likely to be some friction. This is true of the school room, home, or playground. With a good driver these things are held to a minimum. Even if boys fight all the way home on the bus they have less than half as much time to fight as they would have walking home from school where they sometimes fight from the close of school until dark.
4. Transportation is expensive; so is anything else that is worthwhile. Oftentimes enough can be saved on teachers' salaries in consolidated schools (fewer are required than under the small school system) to pay or nearly pay for the transportation.
5. With a definite schedule for the bus there is no more reason for children missing the bus than for them being late to school.
6. Children should wait at some home on the road or else a "waiting station" should be built for them.

7. Pupils can be better cared for at a consolidated school, where there is a bus to take them home, than at a school where they walk and there is nothing to take them home, if they become sick.

8. There are some roads that buses can't get over every day in the year. But good roads usually follow consolidations. We never get roads until we need them. Even if the buses do miss a few days, the children will usually get more from the good school than they would by attending every day in the poorer school.

9. Transportation often tends to increase attendance at school. A report from a principal of a consolidated school during one of the coldest months last winter (Center Point Consolidated School is typical of the results:)

| | Percent of Attendance | | |
|---|---|---|---|
| | Primary | Intermediate | High School |
| Transported pupils .... | 77% | 83% | 84% |
| Town pupils ..... | 57% | 73% | 78% |

"Busing Protests Are Mostly Minor." Ann Arbor News, Ann Arbor, Michigan, September 9, 1971.

This article summarizes reactions to busing across the country a few days after the opening of the school year. In general, court-ordered integration-by-busing plans seemed to be gaining headway in some Northern and Southern schools, although there were instances of parent resistance in others.

In Pontiac, Michigan, busing proceeded, but high tensions still prevailed. Several persons were arrested for disorderly conduct and some students were slightly injured in a racial dispute at a junior high school.

Parents in Boston resisted a transfer-by-busing program and attempted to register their children at their former schools. However, school officials remained firm in their support of the program.

Several thousand students were bused without incident to schools in Mobile, Alabama, but in Birmingham, Alabama, 500 marchers protested court-ordered desegregation.

In Florida, integration was proceeding smoothly with less attendance at specially set up racially segregated private schools than expected. Some difficulties were reported with busing programs in Alexandria, Virginia; Chattanooga, Tennessee; and San Francisco, California.

"Busing: The Court Rules." Newsweek, May 3, 1971, pp. 26-28.

Despite increased informal opposition to busing plans for achieving racially balanced schools, the Supreme Court took a strong 9-0 stand last week and ruled that busing could be ordered to eliminate "all vestiges of state-imposed segregation." To this end the Court approved the practices of busing students and redrawing school district boundaries.

The brunt of the decision fell on Southern cities with their past history of de jure segregation since the wording of the decision seemed definitely aimed at school systems which had been deliberately constructed and maintained to enforce racial segregation.

However, Southerners pointed out that, by some measures, school segregation in the North is already more extensive than it is in the South. Some judges agreed and indicated that they felt that the Court had given them the order to eliminate de facto as well as de jure segregation. As a consequence, Northern school districts feel that they now may be facing the possibility of court action.

Carrigan, Patricia M. "School Desegregation Via Compulsory Pupil Transfer: Early Effects on Elementary School Children." (Report to U.S. Office of Education), September 1969.

This report summarizes the results of a pupil transfer program carried out in Ann Arbor, Michigan public schools in order to achieve a racial balance among the students there. The three pupil populations studied (pupils transferred, pupils in receiving schools, nontransfer pupils in control schools) were, for the most part, midwest-born children of native-born American parents who have resided in the state of Michigan for at least 15 years or so, on the average. The transfer group was predominantly black, the receiving school group was predominantly white, and the nontransfer group had roughly a 50-50 racial split.

The conclusions drawn from a pre- and post-assessment of the population were: While the transfer pupils did in fact earn higher post-transfer scores on the three academic performance measures employed (scholastic aptitude or IQ, reading, and arithmetic), their post-transfer gains were smaller, on the whole, than the gains made by non-transfer and receiving-school pupils.

Self-Esteem, as measured in the present study, appears to have been little affected by a year of desegregated schooling.

Motivation, as measured by two self-report scales, was initially lowest in the transfer group and showed little overall change in the post-transfer year.

Personal Aspiration, as reflected in expressed occupational choice, tended to be lower initially in the transfer group but did not differ significantly in the three populations at the end of the post-transfer year.

Interpersonal Relationships, assessed sociometrically and via self-report measures, tended to show both positive and negative post-transfer effects. On the whole, the transfer pupils appear to have been slightly less well accepted by their receiving-school classmates than they were the previous year by classmates in the de facto segregated school.

Reactions to School, as expressed in a variety of self-report measures, indicated that the transfer pupils found the receiving-school climate less supportive than did children native to the receiving schools, but more supportive than the general climate of the de facto segregated school.

Behavior in School, as described by teachers, appeared to be characterized in the transfer group by generally poorer classroom adjustment and a higher incidence of problems, both before and after the transfer.

Reactions to the Transfer, on the part of the transfer pupils and their parents, were positive in a majority of cases.

The alteration of established behavior patterns and attitudes is not a rapid process; and the present study, dealing with a single year of desegregated schooling, can offer only the most tentative kinds of evidence about progress toward long-range goals.

Cote, William. "Schools Try a Variety of Approaches to Integration." Ann Arbor News, Ann Arbor, Michigan; January 25, 1971, p. 2.

This article summarizes the progress made toward integrated schools in ten of Michigan's districts. Each approached the problem differently.

In Ann Arbor a predominantly black elementary school was closed and its pupils reassigned to six white schools. A study of this program, however, concluded that more than just physical transplanting of a child is required to solve schooling problems.

Kalamazoo redistricted their school system to allow for racially balanced senior high schools.

Lansing has closed schools with predominantly minority group pupils at the elementary and junior high levels and increased the number of minority group students being bused to predominantly white schools.

Battle Creek started desegragating eight white elementary schools in 1968, added Negro history and other courses to the curriculum, and sharply increased the number of black teachers in the system.

Flint's school board redrew distinct boundaries in 1967, increased the hiring of black and chicano teachers, and increased the use of materials about minority groups. New high schools were planned so as to draw students from all parts of the city.

Ferndale and Pontiac are responding to legal charges of intentionally maintaining a racially segregated school and intentionally perpetuating segregation throughout the school system, respectively.

Jackson built a second city high school to facilitate desegration; however, some difficulties have been experienced. Recently the school board adopted a resolution dormally dedicating the district "to the goal of integrating all schools..."

Grand Rapids adopted a master plan calling for the integration of all schools by 1976. However, racial tensions, opposition to busing, and a turnover in school board membership have led to widespread dissatisfaction with the plan.

Detroit has received much publicity concerning its plan to decentralize the school system. Integration was attempted on a voluntary basis through a "magnet" plan. This is currently being challenged by the NAACP.

Gardner, Burleigh B. and others. "The Effect of Busing Black Ghetto Chidren into White Suburban Schools." Social Research, Inc., Chicago, Illinois, July 1970. (ERIC)

This study is an attempt to discover what happens to students' attitudes as a result of experience in desegregated schools. In particular, the students studied were involved in a special busing program called Operation Hospitality, which was carried out by the Chicago Catholic School Board. Through this program, black grade school students from inner city parochial schools are bused to all-white schools, mostly in the suburbs. Although the program has been underway since 1967, it was decided to try to make attitudinal comparisons between bused and non-bused black students and white classmates and non-classmates to see if there is any reasonable evidence of change. In addition, the perfo.mance of students, in terms of grades and achievement tests, is considered. Data for description of students' attitudes were collected by means of a self-administered questionnaire, which included a 20 word-pair semantic differential and a "draw-a-picture" of a person. The report concludes that there is a positive change in the whites' view of blacks; also, the blacks' view of whites is changed in the same direction. Thus, each group improved its image of the other in ways that increase acceptability. Both groups became slightly more interesting to the other. In terms of scholastic performance, there is no significant difference in either grades or performance on tests between the bused and non-bused blacks.

Jensen, Arthur R. "Parent and Teacher Attitudes Toward Integration and Busing." Research Resume, No. 43. California Teachers Association Burlingame, May 1970. (ERIC)

Questionnaires concerning opinions regarding racial integration, busing and ability grouping were solicited in Spring, 1968, just prior to total desegregation of the Berkeley schools, from 337 Berkeley, California elementary school teachers and from the parents of over 8,000 elementary school pupils. Analyses of the results indicated that: (1) the vast majority of Berkeley teachers favored integration and busing, and held attitudes favorable to the school administration's official policies in this area; (2) older teachers were less favorably disposed toward busing and integration than younger teachers; (3) the majority of parents favored integration but approved busing less as a means of achieving integrated schools; (4) more females than males favored busing; (5) there were significant racial differences in opinions on busing; (6) a majority of all racial groups favored ability grouping; (7) favorable attitudes toward busing decreased with the number of years residence in Berkeley; (8) home owners approved of busing less than renters; and (9) favorableness toward integration and busing was positively related to parents' educational level. Teacher and parent opinion questionnaire forms with respective percent responses according to criteria of race, sex, age group, and length of residence in Berkeley are appended. (ERIC)

Laurent, James A. Do Pupil Race and/or School Racial Balance Affect Academic Performance? Bulletin; December 1970. (ERIC)

The purpose of this study is to discover the effects of pupil race and racial balance in schools on educational achievement. A brief review of pertinent background factors and related research literature is presented, followed by an explanation of the study itself. Subjects were one hundred and sixty black and white students in the Tacoma, Washington, public schools, who ranged from primary to junior high school levels. Of the 96 four-group comparisons of variables, 32 at each level, only four showed significant differences. Three of the four were racial effects. Caucasians scoring higher than Negroes in primary mathematics, intermediate language arts, and intermediate composite. No school effects were indicated at any level. The sole interaction effect indicated that Caucasians in nonsegregated schools scored higher than Caucasians in de facto segregated schools on the primary composite. Results of the study suggest that neither pupil race nor racial composition of school per se, considered alone or interactively, seemed to have a substantial effect on academic performance when other relevant variables were controlled. (ERIC)

Mills, Nicolaus. "Busing: Who's Being Taken for a Ride?" Commonweal, Vol. XCVI, No. 3, March 21, 1972.

History of Busing: Its origins not only go back to a time long before desegregation, but even before there were buses. Since at least 1869, public funds have been used to employ vehicles to transport children to and from school. By 1919 pupil transportation at public expense was legal in all 48 states. This was a result of compulsory attendance laws and consolidation of school centers in rural areas. The development of automobiles and improvement of roads increased the demand for pupil transportation after this time.

Extent of Busing: The number of children bused at the time this article was written exceeded 19 million (42.2 percent) making busing the greatest single transportation system in the country. Eighty percent of the school districts in the country maintain one or more vehicles for pupil transportation, with the average being 15.

Who Wants and Benefits?: Traditionally, the greatest demand for busing has come from rural areas, where population is scattered and the consolidated school district is typical. However, it has become an increasingly important factor in urban areas, in response to the trend towards specialized schools. It is also viewed as a solution to overcrowding. Children suffering from handicaps are also frequently bused.

Myths and Realities: With these observations in mind, it is possible to put some of the myths surrounding the current busing issue into perspective.

(1) Busing goes against tradition and represents a break with past approaches to improving education: Busing is an old "tradition," long regarded as a way of equalizing educational opportunity.

(2) Busing is the exception and the neighborhood school is always the most desirable: With the number of school children who are bused as high as it is, it is not possible to regard busing as unusual.

(3) The decision to bus has, until recently, not been guided by social beliefs or principles: Busing in rural areas was guided by the belief that rural students were entitled to as good an education as were urban students. Busing was actually used in the South to maintain a dual school system, by busing whites to one school and blacks, from the same neighborhood, to another.

(4) Riding on the school bus is bad for children: Although there are times when busing might place a hardship on students, children generally see it as a pleasant part of the school day. This point is rarely raised except when busing involves desegregation.

(5) Busing is invariably a financial burden on a community: Busing represents a savings for a community which has supported a dual school system. It is generally cheaper than compensatory education programs.

Petroni, Frank A., Hirsch, Ernest A., and Petroni, C. Lillian. Two, Four, Six, Eight, When You Gonna Integrate? New York: Behavioral Publications, 1970.

This book examines the social-psychological impact of attending an integrated high school. Case studies are utilized involving student subgroups: the black 'elite,' the black militants, white racists and conservatives, the hippies, and the peaceniks.

The range of topics covered includes:

1. Discrimination by teachers, counsellors, and school administrators.
2. The generation gap in attitudes toward racial mixing.
3. Differences in the pattern of racial cleavage in elementary, junior, and senior high school.
4. Areas in which students from minority groups are most and least accepted.
5. The extent of 'true integration'(i.e., other than simply the geographical integration involved in attending the same class).

The authors did not draw definitive conclusions from their interviews, but they did have some final comments and observations:

1. There is not only a similarity between the ways in which blacks, whites, and Mexican-Americans perceive certain events, there are also differences in the manner in which they interpret their social world (these differences are particularly striking within members of the same race).
2. In general, the students believed that good relationships between the races was an important and desirable goal.
3. The students had a rather shallow conception of 'integration.'
4. Blacks and Mexican-Americans who actively seek a place in their schools are discriminated against by fellow blacks and fellow Mexican-Americans.

Purl, Mabel C. and Dawson, Judith. "A Report on the Achievement of Elementary Pupils in Integrated Schools McAteer Project M9-14." Riverside Unified School District, California, May 1970. (ERIC)

Four years after the inception of a school desegregation busing program, the average reading achievement test scores of the bused pupils showed the same trends as those among the receiving pupils and pupils at unaffected schools. Test scores among all kindergarten pupils showed an upward trend, as well as scores in the first, second, and third grades. The changes which have occurred are held to be probably due more to other factors than to desegregation. Desegregation is considered to be more beneficial for the higher achieving minority pupils than for the lower achiever. There is also said to be a significant correlation between the average achievement of bused and receiving pupils. (ERIC)

Stoff, Sheldon. The Two-Way Street. Indianapolis: David-Steward Publishing Company, 1967.

This study was concerned with all the communities in the southern and border states which underwent initial school desegration in the period of time from September 1960 to October 1968; there were 324 such communities.

A questionnaire with eighty items dealing with various aspects of black/white relations was distributed to NAACP members and chief school administrators in the communities; the rate of reply was 62 percent.

The primary examination of the data consisted of a factor analysis; six strong factors emerged which seemed to be associated with peaceful public school desegregation. They are:

Factor I: Favorable School Leadership

1. The chief school administrator led in the development of a school desegregation plan.
2. The chief school administrator supported the desegregation plan.
3. The school board supported the desegregation.

Factor II: The Opposition

1. There was a lack of active opposition to school desegregation by the governor.
2. There was a lack of active opposition to school desegregation by outside organizations.
3. There was a lack of active opposition to school desegregation by local organizations.
4. There was a lack of opposition to school desegregation by the state education department.
5. There was a lack of opposition to school desegregation by local individuals.

Factor III: The Urban Center

1. Smaller communities (under 50,000 persons) had better success.
2. There were no Negroes on the police force.
3. There were no police on duty in plain clothes.
4. Negro and white schools were comparable.
5. Nearby schools desegregated at the same time.

Factor IV: The Practical Community

1. Desegregation was voluntary (not specifically ordered by a court).
2. There was preparation by the white faculty of white students for nonviolent school desegregation.
3. There were nearby desegregated schools.
4. There was a high employment level among the male, nonwhite population.

Factor V: Active Community Support

1. The white clergy supported school desegregation.
2. Respected white community members supported nonviolent school desegregation.
3. School transportation facilities were desegregated.
4. Local newspapers advocated nonviolent school desegregation.

Factor VI: Residual Support

1. There were desegregated labor unions.
2. The state educational agencies supported school desegregation.

"Study of Achievement in Reading of Pupils Transferred from Schools 15 and 37 to Peripheral Schools to Eliminate Overcrowding, to Abandon an Obsolete School, and to Achieve a More Desirable Racial Balance in City Schools." Buffalo Public Schools, New York, March 1967. (ERIC)

Three hundred and fifty Buffalo public school second graders, most of whom were Negro, were transferred in spring 1966, from an inner-city school to schools 90 percent or more white. Two hundred and ten second graders, mostly Negro, were transferred from an adjacent inner-city school to five other peripheral schools. A comparison was made of the pre- and post-transfer scores on the Paragraph Meaning and Word Meaning sections of the Stanford Achievement Test obtained from a sample of 54 transfer students and 60 students not transferred and therefore still in an inner-city school. Pupils who were transferred from the inner-city school to peripheral schools show significantly greater gains when they are compared with pupils who remained in inner-city schools. Further, teacher evaluation of achievement and adjustment to school and classroom procedures on the part of all pupils, grades one through six, who were bussed from the inner-city, show significant positive change. (ERIC)

Sullivan, Neil V. and Steward, Evelyn S. Now is the Time: Integration in the Berkeley Schools. Bloomington, Indiana University Press, 1969 (ERIC)

This book is about the fight for integration in Berkeley, California. Integration in this community was accomplished by busing white children from affluent neighborhoods to black ghetto schools and vice versa. Author recounts the struggles with a traditional, status quo school board, community resistance, and prejudice. Ancedotal stories of the problems the children faced and the solutions they found in the new integrated situations are included. Hailed a success, the Berkeley experiment incorporated a rewritten curriculum which emphasizes black history for all students. Help centers have been added to the middle schools and counselors are readily available to assist all students. Learning laboratories, and the smallest pupil to teacher ratio in California are other aspects of this successful program. (ERIC)

Teele, James E. "The Study of Project Exodus: A School Racial Integration Project in Boston, Massachusetts. Final Report." Harvard School of Public Health; Boston, Massachusetts, November 1969.

This voluntary school integration project uses the open enrollment plan of the Boston School Department in transporting Negro children from predominantly Negro schools in the black district to more racially balanced schools in other parts of Boston. It has involved private financing, intra-city busing, and the initiative and participation of working class ghetto residents. Attitudinal and achievement tests on the children participating in the Project Exodus in grades three to eight were obtained at the following times: the fall of 1967 and the spring of 1968. Similar data were collected on a comparison group of children not enrolled in Exodus and attending predominantly black schools in their neighborhoods. Collection of change data was completed for 151 children. It was found that the children in Exodus showed greater improvement in change in achievement test scores, although they did not show greater fate control. Further data analysis and research are being done to try to more clearly locate the factors related to improvement in both the affective and cognitive areas for Exodus and non-Exodus children. (Appendix A, "Family Experiences in Operation Exodus, the Busing of Negro Children," by James Teele, Ellen Jackson, and Clara Mayo--Community Mental Health Journal Monograph No. 3--is not available from EDRS, but may be obtained from Community Mental Health Journal, Columbia University Press, 605 West 115th Street, New York, New York 10025, for $1.75) (ERIC)

Walberg, Herbert J. "An Evaluation of an Urban-Suburban School Busing Program: Student Achievement and Perception of Class Learning Environments." Paper presented at the Annual Meeting of the American Educational Research Association, New York, New York, February 1971. (ERIC)

This report is part of the evaluation of a program-Metropolitan Council on Educational Opportunity (METCO)--for the academic year 1968-69, which provides screening, placement, and busing services for Negro children from predominantly Negro schools in Boston to predominantly white schools in the surrounding suburbs. In this evaluation, METCO and non-random control children were tested close to the beginning and the end of the academic year, a unique feature of this evaluation being the use of siblings of the bused children as the control group; each control child selected was matched as closely as possible on age. The total population included 737 children in the METCO group and 352 in the control group. With the exception that METCO children gained significantly less than the siblings on mathematics achievement at grade 5-6, there are no significant differences in performance between the two groups from grade 2-12. On a measure of the social environment of the classroom given at grades 3-4 and 5-6, the METCO children perceived their classes as more satisfying; METCO children in grades 5-6 also saw their classes as less difficult and competitive, and as having less friction. The evaluation concludes that school busing programs are a small step in the right direction, but may be doing too little too late. (ERIC)

Wilson, Alan B. "The Consequences of Segregation: Academic Achievement in a Northern Community." California University, Berkeley. Survey Research Center, March 1969.

This monograph, a revision of a report prepared pursuant to a contract with the Commission on Civil Rights, discusses neighborhoods and schools, primary school variation in cognitive development, father absence and school achievement, neighborhood and school segregation, later effects of school segregation, self-concept, aspirations, and behavioral deviance. Questionnaire data were collected from stratified random samples of 17,000 students attending 11 public junior and senior high schools in Western Contra Costa County in the spring of 1965. The final sample comprised 4,077 students. Major substantive findings of the study were as follows: variations in elementary school context make a substantial and significant difference in subsequent academic success at higher grade levels; socio-economic and racial characteristics of students' age-mates have no effect independently on academic achievement of students attending similar schools; social class composition of a school affects the academic development of both Negro and white students in racially segregated or integrated situations; and, given similar social class compositions, racial balance of a school has slight bearing on the academic performance of students. (ERIC)

ELEMENTARY SCHOOL INFORMATION - FAIR CITY

| INDEX | NAME | ADDRESS & TELEPHONE | OPENED | ADDITION | INFORMATION FOR 9-72 | | | | | | PER PUPIL COST |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Teaching Stations | Classrooms in Use | Portables | Enrollment | Teachers | Site (Acres) | |
| 1 | L.J. Horner | 2670 Tyler St. 936-3566 | 1957 | | 12 | 11 | 0 | 367 | 12 | 10 | $605 |
| 2 | P.P. Pumpkineater | 2560 Johnson Ave. 158-8565 | 1941 | 1960 | 27 | 26 | 2 | 808 | 27 | 11 | $612 |
| 3 | Knaveofhearts | 1608 Douglas St. 693-0266 | 1953 | 1965 | 21 | 20 | 0 | 630 | 21 | 9 | $625 |
| 4 | Jackandjill | 600 Hamilton St. 293-0622 | 1923 | | 25 | 24 | 1 | 786 | 25 | 8 | $597 |
| 5 | H.D. Dock | 2775 Burr Drive 168-2791 | 1961 | | 21 | 20 | 0 | 651 | 21 | 9 | $618 |
| 6 | M. Goose | 1414 Jay Avenue 608-4686 | | | 22 | 20 | 0 | 660 | 21 | 10 | $621 |
| 7 | Lamb | 4250 Lee Street 904-0471 | 1959 | | 25 | 23 | 1 | 714 | 24 | 11 | $630 |
| 8 | L.M. Muffet | 2935 Grant Avenue 691-4700 | 1967 | | 16 | 14 | 0 | 455 | 15 | 11 | $627 |
| 9 | J. Sprat | 2135 Allen Lane 893-0265 | 1940 | | 21 | 17 | 0 | 457 | 20 | 15 | $682 |
| 10 | Calicocat | 3540 Darwin Ct. 719-2663 | 1955 | | 28 | 25 | 1 | 772 | 27 | 17 | $701 |
| 11 | T. Blindmice | 800 Lafayette Dr. 888-6866 | 1970 | | 25 | 22 | 0 | 666 | 24 | 18 | $710 |
| 12 | H.D. Diddle | 825 Patton Ave. 821-6597 | 1922 | 1951 | 22 | 20 | 0 | 621 | 22 | 14 | $699 |
| 13 | O.K. Cole | 3800 Stowe St. 895-7771 | 1960 | | 26 | 24 | 1 | 762 | 25 | 12 | $660 |

(Continued)

| | | | | | INFORMATION FOR 9-72 | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| INDEX | NAME | ADDRESS & TELEPHONE | OPENED | ADDITION | Teaching Stations | Classrooms in Use | Portables | Enrollment | Teachers | Site (Acres) | PER PUPIL COST |
| 14 | Shaftoe | 344 Anthony Ave. 650-4665 | 1969 | | 27 | 24 | 1 | 766 | 26 | 15 | $671 |
| 15 | M.M. Contrary | 2250 Keyes Ct. 564-8760 | 1958 | | 19 | 15 | 0 | 339 | 18 | 18 | $705 |
| 16 | Inashoe | 920 Frost St. 561-6623 | 1965 | | 22 | 20 | 0 | 662 | 21 | 13 | $640 |
| 17 | S. Simon | 4377 Blackwell Ave. 112-3434 | 1956 | 1963 | 18 | 17 | 0 | 616 | 18 | 10 | $599 |
| 18 | Hubbard | 3550 King Drive 179-4150 | 1944 | 1952 | 16 | 15 | 0 | 554 | 16 | 11 | $590 |
| 19 | Sixpence | 2775 Wilson St. 178-0200 | 1952 | | 11 | 11 | 0 | 420 | 11 | 9 | $593 |
| 20 | B.B.B. Sheep | 2100 Harding St. 194-0620 | 1925 | | 13 | 12 | 0 | 441 | 12 | 7 | $580 |
| 21 | L.B. Blue | 2543 Lowell Rd. 197-3774 | 1962 | | 8 | 7 | 0 | 241 | 7 | 8 | $586 |

Norman, R.D. and Daley, M.F. "The Comparative Personality Adjustment of Superior and Inferior Readers." Journal of Educational Psychology, 1959, 50(1), 31-36.

The research literature on reading, reports equivocal results about the relationship between reading ability and personality adjustment. The authors feel that the lack of clear results in the past was due to the inadequacy of statistical methods which could handle multivariate and interactional analysis. Thus, the experiment will take advantage of the newer statistical techniques.

The experiment was conducted in a middle class town. The children used in the study were 6th grade boys with I.Q.'s between 84 and 116 as shown by the California Test of Mental Maturity. Two groups were compared on the California Test of Personality within this group of subjects; a superior reading group which had a mean reading level score of grade 8.1 (N=42) was compared with a poor reading group with a mean grade level reading of 3.9 (N=41). An Anova revealed a significant F ratio on the overall adjustment of the two groups, although the differences between groups on subtests were only quasi-significant. They have hypothesized that several groups of variables seem to be related to poor reading ability although no factor analysis was done. These are poor family interaction, frustration-aggression by others, conflicts over dominance, environmental deprivation, aggressive toward others, impulsivity, rejection of others and inferiority feelings.

Raph, J.; Lieberman, D.; and Pascale, P. Influence of a Piaget-Oriented Curriculum on Intellectual Functioning of Lower-Class Kindergarten Children. American Educational Research Association, New York, 1971.

"Intellectual functioning of a lower-class kindergarten class exposed to the introduction of a Piaget-oriented curriculum was compared with intellectual functioning of one comparison lower-class and one comparison middle-class group of children taught in a traditional, activity-centered kindergarten program. Change scores between the beginning and end of a school year were significantly higher for the experimental lower-class group on one standardized measure and on several Piagetian tasks. Main features of the Piaget-oriented curriculum and implications of the results are discussed."

"What is interpreted as the strongest evidence that the innovative features of the kindergarten were, indeed, effective, is shown most clearly on the three most difficult tasks in the Piagetian battery, Reverse Linear Ordering, Seriation II, and Conservation. In these instances, the experimental group made significant gains, the comparison lower- and middle-class groups did not. This outcome may be said to reflect something of Piaget's position that pedagogical interventions can accelerate or complete spontaneous development, but cannot change the order of constructions (Piaget, 1970)."

California Test of Mental Maturity (Short Form) (1 Form)

Purpose: To provide information about the functional capacities that are basic to learning, problem-solving, and responding to new situations.

Type of Test: Group intelligence

Range: Grades K-16

Time for Administration: 34-48 minutes

Cost: $4.20 per 35 tests, $.10 per scoring key

Subtests:
- Opposites
- Similarities
- Analogies
- Delayed Recall
- Numerical Values
- Number Problems
- Verbal Comprehension

Levels:
- 0 - Preprimary
- 1 - Primary
- 1H - Spans range between grades 3-4
- 2 - Elementary
- 2H - Spans range between grades 6-7
- 3 - Junior High
- 4 - Secondary
- 5 - Advanced

Scores: Scaled; Language IQ, Non-Language IQ, Total IQ

Nature of Test:

The California Short-Form of Mental Maturity consists of seven administrative test units, measuring aspects of mental ability, which contribute to a pattern of summary and derived scores (mental ages, intelligence quotients, standard scores, stanines, and percentile ranks) that are interpreted within a framework of inter and intra-individual differences. In addition to assessing the development of an individual or group with reference to national performance standards at each age level, the test results provide data as to the nature and potential of the abilities possessed by the individual. The Short-Form thus serves both survey and analytical purposes for educators, counselors, psychologists, and employers in a wide variety of testing situations.

California Short Form Test of Mental Maturity - Mean Intelligence Quotients & Standard Deviations of Mental Maturity Series, 1963 Revision & the Stanford-Binet Intelligence Scale Form LM 1960 Revision, Based on the Performance of the Sample Groups

| Short-Form Level | Type of CTMM IQ | IQ Means | | IQ SD's | |
|---|---|---|---|---|---|
| | | Cal. | Binet | Cal. | Binet |
| Level 0 | Total | 109.4 | 108.2 | 14.9 | 14.8 |
| | Language | 109.4 | | 16.6 | |
| | Non-Language | 107.6 | | 14. | |
| Level 1 | | Cal. | Binet | Cal. | Binet |
| | Total | 110.2 | 109.2 | 15.1 | 15.0 |
| | Language | 109.3 | | 15.5 | |
| | Non-Language | 109.8 | | 15.5 | |

| Short-Form Level | Type of CTMM IQ | IQ Means | | IQ S.D.'s | |
|---|---|---|---|---|---|
| Level 2 | | Cal. | Binet | Cal. | Binet |
| | Total | 103.7 | 106.7 | 14.2 | 13.8 |
| | Language | 104.5 | | 15.0 | |
| | Non-Language | 102.3 | | 14.2 | |
| Level 3 | | Cal. | Binet | Cal. | Binet |
| | Total | 105.5 | 106.5 | 15.0 | 14.9 |
| | Language | 104.9 | | 14.8 | |
| | Non-Language | 105.0 | | 15.5 | |
| Level 4 | | Cal. | Binet | Cal. | Binet |
| | Total | 108.9 | 111.8 | 13.9 | 14.3 |
| | Language | 106.1 | | 14.5 | |
| | Non-Language | 109.2 | | 14.0 | |
| Level 5 | | Cal. | Binet | Cal. | Binet |
| | Total | 113.9 | 116.4 | 14.5 | 14.9 |
| | Language | 112.1 | | 14.5 | |
| | Non-Language | 111.9 | | 14.1 | |
| All Levels | | Cal. | Binet | Cal. | Binet |
| | Total | 108.6 | 109.7 | 15.0 | 15.0 |
| | Language | 107.8 | | 15.4 | |
| | Non-Language | 107.6 | | 15.0 | |

California Test of Mental Maturity (Short Form)

| Variables | Reliability | Validity Coefficients of Correlation |
|---|---|---|
| Level 0 | Short Form | |
| I.Q. Total | .78 | .68 |
| Language | .70 | .61 |
| Non-Language | .59 | .56 |
| M.A. Total | .78 | .80 |
| Language | .70 | .74 |
| Non-Language | .59 | .71 |
| Level 1 | | |
| I.Q. Total | .83 | .70 |
| Language | .73 | .64 |
| Non-Language | .72 | .57 |
| M.A. Total | .83 | .89 |
| Language | .73 | .87 |
| Non-Language | .72 | .83 |
| Level 2 | | |
| I.Q. Total | .95 | .74 |
| Language | .94 | .70 |
| Non-Language | .89 | .61 |
| M.A. Total | .95 | .80 |
| Language | .94 | .76 |
| Non-Language | .89 | .69 |

| Variables | Reliability Short Form | Validity Coefficients of Correlation |
|---|---|---|
| Level 0 | | |
| Level 3 | | |
| I.Q. Total | .91 | .78 |
| Language | .90 | .77 |
| Non-Language | .81 | .65 |
| M.A. Total | .91 | .78 |
| Language | .90 | .77 |
| Non-Language | .81 | .63 |
| Level 4 | | |
| I.Q. Total | .92 | .71 |
| Language | .91 | .68 |
| Non-Language | .83 | .57 |
| M.A. Total | .92 | .76 |
| Language | .91 | .73 |
| Non-Language | .83 | .63 |
| Level 5 | | |
| I.Q. Total | .91 | .75 |
| Language | .90 | .75 |
| Non-Language | .82 | .56 |
| M.A. Total | .91 | .75 |
| Language | .90 | .75 |
| Non-Language | .82 | .57 |
| Combined | | |
| I.Q. Total | .91 | .73 |
| Language | .90 | .69 |
| Non-Language | .82 | .60 |
| M.A. Total | .91 | .96 |
| Language | .90 | .96 |
| Non-Language | .82 | .94 |

California Test of Mental Maturity (Critique)

See Also:

Anderson, H.E. and Leton, D.A. "Factor Analysis of the California Test of Mental Maturity." Educational and Psychological Measurements, 1964, 24(3), 513-523.

Norman, R.D. and Daley, M.F. "The Comparative Personality Adjustment of Superior and Inferior Readers." Journal of Educational Psychology, 1959, 50(1) 31-36.

California Test of Personality (2 Forms)

Purpose: To identify and reveal the status of certain highly important factors in personality and social adjustment usually designated as 'intangibles.'

Type of Test: Group personality

Range: Grades 1-adult (may also be used for kindergartin )

Time for Administration: 45-60 minutes

Cost: $3.50 per 35 tests

Subtests:
- Self-Reliance
- Sense of Personal Worth
- Sense of Personal Freedom
- Feeling of Belonging
- Withdrawing Tendencies
- Nervous Symptoms
- Social Standards
- Social Skills
- Anti-Social Tendencies
- Family Relations
- (School/Occupational) Relations
- Community Relations

Scores: Scaled for each subtest:

Personal Adjustment (Sum of first 5 subtests)

Social Adjustment (Sum of last 5 subtests)

Total Adjustment (Sum of personal & social adjustment)

Nature of Test:

The California Test of Personality has been designed to identify and reveal the status of certain highly important factors in personality and social adjustment usually designated as intangibles. These are the factors that defy appraisal or diagnosis by means of ordinary ability and achievement tests. Measurements of capacity, skill, and achievement, important as they are, do not constitute a complete picture of a functioning personality.

California Test of Personality - Means & Standard Deviations of T Scores on California Test of Personality Variables for Both Groups

| Variables | Inferior Readers | | Superior Readers | |
|---|---|---|---|---|
| | Mean | SD | Mean | SD |
| Personal Adjustment: | | | | |
| Self-Reliance | 46.90 | 6.77 | 52.28 | 7.27 |
| Sense of Personal Worth | 50.49 | 8.30 | 57.33 | 7.36 |
| Sense of Personal Freedom | 46.17 | 7.57 | 53.17 | 6.88 |
| Feeling of Belonging | 47.41 | 8.81 | 56.50 | 6.79 |
| Withdrawing Tendencies (Freedom from) | 51.32 | 7.84 | 59.38 | 6.96 |
| Nervous Symptoms (Freedom from) | 49.44 | 11.43 | 55.17 | 7.61 |
| Social Adjustment: | | | | |
| Social Standards | 43.95 | 8.98 | 53.83 | 7.36 |
| Social Skills | 49.17 | 9.84 | 54.36 | 10.19 |
| Anti-Social Tendencies (Freedom from) | 44.80 | 10.21 | 52.14 | 9.44 |
| Family Relations | 45.12 | 8.47 | 52.52 | 6.83 |
| School Relations | 50.05 | 9.79 | 57.71 | 7.14 |
| Community Relations | 43.07 | 8.72 | 50.43 | 5.22 |
| Total Adjustment: | | | | |

California Test of Personality

Reliability:

| Components: | Prim. r | Elem. r | Inter. r |
|---|---|---|---|
| Personal Adjustment | .83 | .93 | .93 |
| A. Self-Reliance | .73 | .64 | .70 |
| B. Sense of Personal Worth | .82 | .79 | .75 |
| C. Sense of Personal Freedom | .73 | .79 | .92 |
| D. Feeling of Belonging | .70 | .77 | .97 |
| E. Withdrawing Tendencies (Fdm.) | .71 | .83 | .83 |
| F. Nervous Symptoms (Fdm.) | .87 | .82 | .82 |
| Social Adjustment | .80 | .92 | .94 |
| A. Social Standards | .51 | .59 | .94 |
| B. Social Skills | .70 | .73 | .75 |
| C. Anti-Social Tendencies (Fdm.) | .82 | .77 | .86 |
| D. Family Relations | .82 | .77 | .92 |
| E. School Relations | .70 | .78 | .86 |
| F. Community Relations | .78 | .79 | .87 |
| Total Adjustment | .88 | .94 | .96 |

Reliability

| Components: | Sec. r | Adult r |
|---|---|---|
| Personal Adjustment | .90 | .93 |
| A. Self-Reliance | .70 | .81 |
| B. Sense of Personal Worth | .77 | .66 |
| C. Sense of Personal Freedom | .84 | .87 |
| D. Feeling of Belonging | .91 | .87 |
| E. Withdrawing Tendencies (Fdm.) | .86 | .86 |
| F. Nervous Symptoms (Fdm.) | .82 | .81 |
| Social Adjustment | .89 | .93 |
| A. Social Standards | .84 | .76 |
| B. Social Skills | .86 | .70 |
| C. Anti-Social Tendencies (Fdm.) | .84 | .75 |
| D. Family Relations | .91 | .91 |
| E. School Relations | .73 | .84 |
| F. Community Relations | .78 | .77 |
| Total Adjustment | .93 | .95 |

| California Test of Personality<br>Validity: | Content & Construct | Concurrent & Predictive |
|---|---|---|
| Personal Adjustment | Very Good | Poor |
| A. Self-Reliance | Good | Poor |
| B. Sense of Personal Worth | Good | Poor |
| C. Sense of Personal Freedom | Good | Poor |
| D. Feeling of Belonging | Good | Poor |
| E. Withdrawing Tendencies | Good | Poor |
| F. Nervous Symptoms | Very Good | Poor |
| Social Adjustment | Very Good | Poor |
| A. Social Standards | Good | Poor |
| B. Social Skills | Good | Poor |
| C. Anti-Social Tendencies | Good | Poor |
| D. Family Relations | Good | Poor |
| E. School Relations | Good | Poor |
| F. Community Relations | Good | Poor |
| Total Adjustment | | |

California Test of Personality (Critique)

See Also:

Norman, R.D. and Daley, M.F. "The Comparative Personality Adjustment of Superior and Inferior Readers." Journal of Educational Psychology, 1959, 50(1), 31-36.

Deutsch SES Index

The index of cultural deprivation (C.D.) includes the following three components:

1. The father's occupation on a 4-point scale (or the mother's occupation if no father is in the home).

1 point - unskilled or unemployed
2 points - semiskilled
3 points - skilled
4 points - professional

2. Number of years of education completed by the parents (an average of the two, or the mother's education only if no father is in the home).

3. Density in the home is defined as the number of rooms divided by the number of people living in the home.

In the number of rooms are included the kitchen and the bathroom. A shared bathroom is counted as half a room.

Occupation and education are given a full weight, but density is given a 1/2 weight. Each component is divided by its standard deviation to equate the variability of all components. In other words, the C.D. rating is the sum of three $z$ scores, except that density is given a 1/2 weight.

The above gives a definitional formula. In actuality, the following computation formula is used:

$$1/2\,(\text{Education}) + 2\,(\text{Occupation}) + 2\,\frac{\text{Rooms}}{\text{People}}$$

Reflected in the computational formula are the following approximate standard deviations from the original Perry School population: 2 for education, 1/2 for occupation, and 1/4 for density.*

An example from an actual case may best clarify the computational formula. If pertinent data are:

Father's occupation: Unskilled factory work (1 point)
Mother's education: 11 years
Father's education: 7 years
Number of rooms: 6
Number of people living in home: 9

$$\text{C.D.} = 1/2(9) + 2(1) + 2\left(\frac{6}{9}\right) = 8.3$$

---

*A score of 10 or below makes the family a low SES family and thus "culturally deprived."

Gates-MacGinitie Reading Tests (3 Forms)

Purpose: To provide an objective measure of how rapidly students can read with understanding, sample the student's reading vocabulary, and measure the student's ability to read complete prose with understanding.

Type of Test: Group reading

Range: Grades 1-12

Time for Administration: 40-46 minutes

Cost: $1.75-$2.75 per 35 tests

Subtests: Vocabulary, Comprehension, Speed and Accuracy

Scores: Raw; one for each subtest

Nature of Test:

The vocabulary test asks the examinee to decide which of the possible answers is most similar in meaning to a given stimulus. The comprehension subtest requires the student to read a sentence or short passage and respond to questions, select the most appropriate words to fit into the blanks in the paragraph. The speed and accuracy subtest contains short paragraphs of uniform difficulty followed by a question or an incomplete statement each with a choice to be made from four possible words.

Gates-MacGinitie Reading Tests

| Test | Grade | Subtest Score | Average Raw Score Mean | Average Raw Score S.D. | Alternate Form Reliability |
|---|---|---|---|---|---|
| Primary A | 1 | Vocabulary | 20.7 | 9.8 | .86 |
| | | Comprehension | 11.4 | 6.1 | .83 |
| Primary B | 2 | Vocabulary | 27.7 | 10.1 | .87 |
| | | Comprehension | 17.2 | 7.8 | .81 |
| Primary C | 3 | Vocabulary | 32.9 | 8.3 | .85 |
| | | Comprehension | 25.0 | 10.3 | .87 |
| Primary CS | 3 | Speed No. Attem. | 15.7 | 6.5 | .72 |
| | | *Speed No.Correc. | 14.1 | 7.1 | .86 |
| Survey D | 4 | Vocabulary | 18.6 | 7.9 | .85 |
| | | Comprehension | 20.3 | 10.8 | .83 |
| | | Speed No. Attem. | 16.8 | 6.6 | .67 |
| | | Speed No.Correc. | 12.0 | 5.0 | .80 |
| | 5 | Vocabulary | 27.2 | 9.3 | .87 |
| | | Comprehension | 32.2 | 11.9 | .89 |
| | | Speed No. Attem. | 21.6 | 4.4 | .75 |
| | | Speed No.Correc. | 17.4 | 4.8 | .76 |
| | 6 | Vocabulary | 31.3 | 8.3 | .85 |
| | | Comprehension | 36.9 | 11.1 | .87 |
| | | Speed No. Attem. | 24.8 | 3.3 | .72 |
| | | Speed No.Correc. | 17.9 | 5.7 | .78 |
| Survey E | 7 | Vocabulary | 21.1 | 6.3 | .78 |
| | | Comprehension | 34.9 | 7.0 | .81 |
| | | Speed No. Attem. | 13.1 | 4.7 | .69 |
| | | Speed No.Correc. | 11.0 | 2.5 | .70 |

*Accuracy

Gates-MacGinitie Reading Tests

| Test | Grade | Subtest Score | Average Raw Score Mean | Average Raw Score S.D. | Alternate Form Reliability |
|---|---|---|---|---|---|
| | 8 | Vocabulary | 26.0 | 7.2 | .80 |
| | | Comprehension | 33.1 | 9.4 | .81 |
| | | Speed No. Attem. | 14.3 | 2.9 | .72 |
| | | Speed No.Correc. | 13.1 | 4.3 | .76 |
| | 9 | Vocabulary | 23.0 | 7.3 | .83 |
| | | Comprehension | 30.1 | 8.8 | .80 |
| | | Speed No. Attem. | 16.1 | 3.2 | .68 |
| | | Speed No.Correc. | 15.2 | 3.3 | .77 |
| Survey F | 10 | Vocabulary | 27.9 | 9.5 | .90 |
| | | Comprehension | 29.0 | 10.4 | .91 |
| | | Speed No. Attem. | 18.2 | 6.1 | .73 |
| | | Speed No.Correc. | 16.0 | 5.5 | .78 |
| | 11 | Vocabulary | 29.9 | 9.4 | .92 |
| | | Comprehension | 31.0 | 10.3 | .88 |
| | | Speed No. Attem. | 19.0 | 5.9 | .64 |
| | | Speed No.Correc. | 17.1 | 5.7 | .81 |
| | 12 | Vocabulary | 31.8 | 9.7 | .88 |
| | | Comprehension | 33.3 | 10.2 | .85 |
| | | Speed No. Attem. | 19.7 | 6.2 | .78 |
| | | Speed No.Correc. | 17.9 | 5.8 | .80 |

Gates-MacGinitie Reading Tests

Reliability:

| Subtest | Split Half | Alternate Form |
|---|---|---|
| Comprehension | .89-.96 | .80-.89 |
| Vocabulary | .88-.93 | .78-.87 |
| Speed | | .67-.75 |
| Accuracy | | .70-.86 |

Validity: Content and Construct - Good
Concurrent and Predictive - Poor

Gates-MacGinitie Reading Tests (Critique)

See Also:

Powell, W.R. "Gates-MacGinitie Reading Tests." Journal of Educational Measurement, 1969, 6(2), 114-116.

Stanford Achievement Tests (3 Forms)

Purpose: To test the educational achievements that are commonly expected of students in a modern comprehensive school

Type of Test: Individual or group achievement

Range: Grades 1-12

Time for Administration: Varies from 127-440 minutes (4-7 sessions)

Cost: $5.65-$9.75 per 35 tests, $.60-$1.60 per set of keys

Subtests:

Primary I Battery (Grades 1.5-2.5)

| | |
|---|---|
| Word Reading (Word Meaning) | Spelling |
| Paragraph Meaning | Word Study Skills |
| Vocabulary | Arithmetic (Concepts) |

Primary II Battery (Grades 2.5-3.9)

| | |
|---|---|
| Word Meaning | Spelling |
| Paragraph Meaning | Word Study Skills |
| Science & Social Studies Concepts | Language |
| Arithmetic Concepts | Arithmetic Computation |

Intermediate I Battery (Grades 4.0-5.5)

| | |
|---|---|
| Word Meaning | Social Studies |
| Paragraph Meaning | Arithmetic Computation |
| Spelling | Arithmetic Concepts |
| Word Study Skills | Arithmetic Applications |
| Language | Science |

Intermediate II Battery (Grades 5.5-6.9)

| | |
|---|---|
| Word Meaning | Arithmetic Computation |
| Paragraph Meaning | Arithmetic Concepts |
| Spelling | Arithmetic Applications |
| Language | Social Studies |
| Science | |

Advanced Battery (Grades 7.0-9.9)

| | |
|---|---|
| Paragraph Meaning | Arithmetic Concepts |
| Spelling | Arithmetic Applications |
| Language | Social Studies |
| Arithmetic Computation | Science |

High School Battery (Grades 9-12)

| | |
|---|---|
| Reading | Mathematics |
| Spelling | Social Studies |
| English | Arts & Humanities |
| Numerical Competence | Business & Economics |
| Science | Technical Comprehension |

Scores: Raw; for each subtest

Stanford Achievement Tests

| | | | | | | |
|---|---|---|---|---|---|---|
| Subtest: | Word Reading | Grade 1 | Means | 16.1 | S.D. | 6.88 |
| | Paragraph Meaning | Grade 1 | Means | 14.1 | S.D. | 7.05 |
| | Vocabulary | Grade 1 | Means | 18.8 | S.D. | 6.69 |
| | Spelling | Grade 1 | Means | 7.7 | S.D. | 6.05 |
| | Word Study Skills | Grade 1 | Means | 31.7 | S.D. | 9.47 |
| | Arithmetic | Grade 1 | Means | 30.4 | S.D. | 14.43 |
| Subtest: | Word Meaning | Grade 2 | Means | 16.0 | S.D. | 6.3 |
| | Paragraph Meaning | Grade 2 | Means | 28.1 | S.D. | 12.35 |
| | Sci. & Social Studies Concepts | Grade 2 | Means | 17.6 | S.D. | 4.97 |
| | Spelling | Grade 2 | Means | 10.9 | S.D. | 7.24 |
| | Word Study Skills | Grade 2 | Means | 34.5 | S.D. | 12.90 |
| | Language | Grade 2 | Means | 35.3 | S.D. | 9.72 |
| | Arith. Computation | Grade 2 | Means | 18.4 | S.D. | 8.38 |
| | Arith. Concepts | Grade 2 | Means | 17.2 | S.D. | 7.42 |
| Subtest: | Word Meaning | Grade 3 | Means | 23.5 | S.D. | 6.68 |
| | Paragraph Meaning | Grade 3 | Means | 40.1 | S.D. | 11.33 |
| | Sci. & Social Studies Concepts | Grade 3 | Means | 22.1 | S.D. | 5.70 |
| | Spelling | Grade 3 | Means | 18.4 | S.D. | 7.73 |
| | Word Study Skills | Grade 3 | Means | 42.1 | S.D. | 13.23 |
| | Language | Grade 3 | Means | 46.4 | S.D. | 11.40 |
| | Arith. Computation | Grade 3 | Means | 32.4 | S.D. | 9.30 |
| | Arith. Concepts | Grade 3 | Means | 26.6 | S.D. | 9.58 |
| Subtest: | Word Meaning | Grade 4 | Means | 19.0 | S.D. | 7.94 |
| | Paragraph Meaning | Grade 4 | Means | 30.6 | S.D. | 11.22 |
| | Spelling | Grade 4 | Means | 28.6 | S.D. | 11.37 |
| | Word Study Skills | Grade 4 | Means | 37.9 | S.D. | 12.92 |
| | Language | Grade 4 | Means | 73.6 | S.D. | 17.15 |
| | Arith. Computation | Grade 4 | Means | 19.2 | S.D. | 7.07 |
| | Arith. Concepts | Grade 4 | Means | 15.4 | S.D. | 6.86 |
| | Arith. Applications | Grade 4 | Means | 16.0 | S.D. | 6.46 |
| | Social Studies | Grade 4 | Means | 25.8 | S.D. | 8.28 |
| | Science | Grade 4 | Means | 31.1 | S.D. | 11.16 |
| Subtest: | Word Meaning | Grade 5 | Means | 22.3 | S.D. | 8.76 |
| | Paragraph Meaning | Grade 5 | Means | 31.7 | S.D. | 11.96 |
| | Spelling | Grade 5 | Means | 26.8 | S.D. | 10.14 |
| | Language | Grade 5 | Means | 81.2 | S.D. | 18.72 |
| | Arith. Computation | Grade 5 | Means | 14.8 | S.D. | 5.18 |
| | Arith. Concepts | Grade 5 | Means | 12.1 | S.D. | 5.15 |
| | Arith. Applications | Grade 5 | Means | 16.6 | S.D. | 6.05 |
| | Social Studies | Grade 5 | Means | 37.3 | S.D. | 11.79 |
| | Science | Grade 5 | Means | 30.0 | S.D. | 9.53 |

(Continued)

Stanford Achievement Tests

| | | | | | | |
|---|---|---|---|---|---|---|
| Subtest: | Word Meaning | Grade 6 | Means | 27.6 | S.D. | 9.25 |
| | Paragraph Meaning | Grade 6 | Means | 38.4 | S.D. | 11.94 |
| | Spelling | Grade 6 | Means | 33.1 | S.D. | 11.26 |
| | Language | Grade 6 | Means | 90.5 | S.D. | 18.88 |
| | Arith. Computation | Grade 6 | Means | 21.7 | S.D. | 7.21 |
| | Arith. Concepts | Grade 6 | Means | 16.8 | S.D. | 6.80 |
| | Arith. Applications | Grade 6 | Means | 21.5 | S.D. | 7.83 |
| | Social Studies | Grade 6 | Means | 44.6 | S.D. | 12.41 |
| | Science | Grade 6 | Means | 35.0 | S.D. | 9.69 |
| Subtest: | Paragraph Meaning | Grade 7 | Means | 31.0 | S.D. | 11.86 |
| | Spelling | Grade 7 | Means | 26.4 | S.D. | 10.88 |
| | Language | Grade 7 | Means | 92.5 | S.D. | 17.71 |
| | Arith. Computation | Grade 7 | Means | 19.2 | S.D. | 7.28 |
| | Arith. Concepts | Grade 7 | Means | 16.5 | S.D. | 6.47 |
| | Arith. Applications | Grade 7 | Means | 13.7 | S.D. | 5.14 |
| | Social Studies | Grade 7 | Means | 44.9 | S.D. | 13.15 |
| | Science | Grade 7 | Means | 32.2 | S.D. | 9.82 |
| Subtest: | Paragraph Meaning | Grade 8 | Means | 35.6 | S.D. | 11.95 |
| | Spelling | Grade 8 | Means | 31.0 | S.D. | 12.23 |
| | Language | Grade 8 | Means | 98.0 | S.D. | 17.70 |
| | Arith. Computation | Grade 8 | Means | 23.5 | S.D. | 8.35 |
| | Arith. Concepts | Grade 8 | Means | 20.4 | S.D. | 7.58 |
| | Arith. Applications | Grade 8 | Means | 16.4 | S.D. | 5.76 |
| | Social Studies | Grade 8 | Means | 50.3 | S.D. | 13.52 |
| | Science | Grade 8 | Means | 35.6 | S.D. | 10.09 |
| Subtest: | Paragraph Meaning | Grade 9 | Means | 38.0 | S.D. | 12.25 |
| | Spelling | Grade 9 | Means | 32.9 | S.D. | 12.50 |
| | Language | Grade 9 | Means | 101.5 | S.D. | 18.12 |
| | Arith. Computation | Grade 9 | Means | 25.8 | S.D. | 8.84 |
| | Arith. Concepts | Grade 9 | Means | 22.7 | S.D. | 8.21 |
| | Arith. Applications | Grade 9 | Means | 17.5 | S.D. | 5.68 |
| | Social Studies | Grade 9 | Means | 52.6 | S.D. | 13.60 |
| | Science | Grade 9 | Means | 38.0 | S.D. | 9.98 |

Stanford Achievement Tests

| Primary I Grade I | $r_{11}$ |
|---|---|
| Word Reading | .85 |
| Paragh. Meaning | .90 |
| Vocabulary | .79 |
| Spelling | .92 |
| Word Study Skills | .88 |
| Arithmetic | .95 |

| Primary II Grade 2 | $r_{11}$ |
|---|---|
| Word Meaning | .85 |
| Paragh. Meaning | .93 |
| Sci. Soc. Studies Concepts | .66 |
| Spelling | .94 |
| Word Study Skills | .92 |
| Language | .84 |
| Arith. Comput. | .93 |
| Arith. Concepts | .86 |

(Continued)

Stanford Achievement Tests

| Primary II Grade 3 | $r_{11}$ |
|---|---|
| Word Meaning | .87 |
| Paragraph Meaning | .93 |
| Sci. Soc. Studies Concepts | .79 |
| Spelling | .94 |
| Word Study Skills | .94 |
| Language | .89 |
| Arith. Comput. | .93 |
| Arith. Concepts | .90 |

| Intermediate I Grade 4 | $r_{11}$ |
|---|---|
| Word Meaning | .90 |
| Paragraph Meaning | .92 |
| Spelling | .93 |
| Word Study Skills | .94 |
| Language | .93 |
| Arith. Comput. | .89 |
| Arith. Concepts | .86 |
| Arith. Applications | .86 |
| Social Studies | .86 |
| Science | .91 |

| Intermediate II Grade 5 | $r_{11}$ |
|---|---|
| Word Meaning | .89 |
| Paragraph Meaning | .93 |
| Spelling | .91 |
| Language | .93 |
| Arith. Comput. | .83 |
| Arith. Concepts | .77 |
| Arith. Applications | .84 |
| Social Studies | .89 |
| Science | .88 |

| Intermediate II Grade 6 | $r_{11}$ |
|---|---|
| Word Meaning | .90 |
| Paragraph Meaning | .93 |
| Spelling | .93 |
| Language | .95 |
| Arith. Comput. | .89 |
| Arith. Concepts | .85 |
| Arith. Applications | .89 |
| Social Studies | .91 |
| Science | .90 |

| Advanced Grade 7 | $r_{11}$ |
|---|---|
| Paragraph Meaning | .93 |
| Spelling | .91 |
| Language | .94 |
| Arith. Comput. | .87 |
| Arith. Concepts | .82 |
| Arith. Applications | .77 |
| Social Studies | .89 |
| Science | .88 |

| Advanced Grade 8 | $r_{11}$ |
|---|---|
| Paragraph Meaning | .93 |
| Spelling | .92 |
| Language | .94 |
| Arith. Comput. | .90 |
| Arith. Concepts | .86 |
| Arith. Applications | .81 |
| Social Studies | .91 |
| Science | .89 |

| Advanced Grade 9 | $r_{11}$ |
|---|---|
| Paragraph Meaning | .94 |
| Spelling | .94 |
| Language | .94 |
| Arith. Comput. | .92 |
| Arith. Concepts | .87 |
| Arith. Applications | .79 |
| Social Studies | .90 |
| Science | .90 |

| Grade 9 (Reliability Coefficients) | $r_{11}$ |
|---|---|
| English | .94 |
| Numerical Competence | .90 |
| Math - Part A | .84 |
| Math - Total | .84 |
| Reading | .90 |
| Science - Part A | .87 |
| Science - Total | .87 |
| Social Studies | .85 |
| Spelling | .89 |
| Arts & Humanities | .91 |
| Business & Economics | .88 |
| Tech. Comprehension | .92 |

(Continued)

Stanford Achievement Tests (Reliability Coefficients)

| Grade 10 | $r_a$ | Grade 11 | $r_a$ |
|---|---|---|---|
| English | .94 | English | .94 |
| Numerical Competence | .90 | Numerical Competence | .92 |
| Math - Part A | .89 | Math - Part A | .92 |
| Math - Total | .89 | Math - Total | .93 |
| Reading | .91 | Reading | .92 |
| Science - Part A | .88 | Science - Part A | .87 |
| Science - Total | .89 | Science - Total | .89 |
| Social Studies | .85 | Social Studies | .88 |
| Spelling | .90 | Spelling | .91 |
| Arts & Humanities | .92 | Arts & Humanities | .93 |
| Business & Economics | .88 | Business & Economics | .90 |
| Tech. Comprehension | .93 | Tech. Comprehension | .94 |

| Grade 12 | $r_a$ |
|---|---|
| English | .94 |
| Numerical Competence | .92 |
| Math - Part A | .92 |
| Math - Total | .94 |
| Reading | .91 |
| Science - Part A | .87 |
| Science - Total | .89 |
| Social Studies | .88 |
| Spelling | .92 |
| Arts & Humanities | .93 |
| Business & Economics | .90 |
| Tech. Comprehension | .94 |

| Stanford Achievement Tests | Validity | |
|---|---|---|
| | Content & Construct | Concurrent & Predictive |
| Arith. Applications | Good | Good |
| Arith. Computation | Very Good | Good |
| Arith. Concepts | Very Good | Fair |
| Arith. Total | Good | Fair |
| Language | Very Good | Fair |
| Paragraph Meaning | Very Good | Fair |
| Vocabulary | Very Good | Poor |
| Word Meaning | Very Good | Fair |
| Science | Good | Good |
| Science & Social Studies Concepts | | Good<br>Poor |
| Social Studies | Good | Good |
| Spelling | Very Good | Fair |
| Word Study Skills | Very Good | Fair |

Stanford Achievement Tests (Critiques)

See Also:

Olson, A.V. and Johnson, C.I. "Structure and Predictive Validity of the Frostig Developmental Test of Visual Perception in Grades One and Three." The Journal of Special Education, 1970, 4(1), 49-52.

Stake, R.E. and Hastings, J.T. "The Stanford Achievement Test Battery." Personnel and Guidance Journal, 1964, 43(10), 178-184.

Brim, O.B. "American Attitudes Toward Intelligence Tests." American Psychologists, 20, 1965, 125-130.

This article is an outgrowth of a project under the Russell Sage Foundation which investigated the social consequences of administering standardized ability tests. The author elaborates the various objections which have been raised about testing. He divides the objections into two categories: superficial or apparent criticisms and basic sources of criticisms. The former criticisms are felt to be ones which could be almost alleviated with the proper controls and education. The latter kind of criticisms are harder to alleviate since these deal with more basic issues such as long-range self-esteem, and the issue of equality of all men. He concludes by saying: "Our review of these four fundamental sources of opposition to tests leads to the conclusion that psychologists must be concerned about the society in which tests are used. Testing does not occur in isolation; there is always a social context. Test scores have a social meaning. They have impact on man's self-esteem; they influence his life chances; they engage his deepest political and social attitudes. These forces must be understood as part of the social setting in which intelligence testing is carried on."

Bryant, N.D. and Della-Piana, M.M. Critique of Gates Reading Diagnostic Tests. In O.K. Buros (ed.) The Sixth Mental Measurement Yearbook. Highland Park, New Jersey: The Gryphon Press, 1965, 824.

The reviewer feels this test has both good and bad aspects. "One of the strengths of this battery is the studding of suggestions for profile analysis throughout the manual. Even the user of other batteries would gain much from these sections." It contains many subtests which are valuable in identifying reading problems, including things like auditory blending, oral reading, phrase recognition, etc. A weakness of the test is the lack of normative samples, the lack of good reliability data, and the lack of reading comprehension subtest. Because of the lack of good normative data, it is possible for the examiner to misinterpret scores on the test. "If this battery is used as intended, the user must interpret norms where the reference group is not entirely clear, needs to make decisions as to whether a test should be given and in what order, and needs to consider alternative interpretations of test scores. Thus, administering and interpreting this battery requires a sophisticated examiner willing to pay the price of considerable training, formal or informal." The casual user, whether it be a test professional or a teacher, should be wary of using this test. For such a user the Durrell seems to be a more reliable indicator of reading difficulty.

Churchill, W.D. and Smith, S.E. "The Relationship of the 1960 Revised Stanford-Binet Intelligence Scale to Intelligence and Achievement Test Scores Over a Three-Year Period." Educational and Psychological Measurement, 1966, 26(4), 1015-1020.

The primary objective of this study was to determine the relationship between the 1960 Stanford-Binet Intelligence Scale (S-B) and Lorge-Thorndike IQ's obtained three years later. A secondary objective was to determine the relationship between the 1960 S-B and achievement as measured by the Iowa Test of Basic Skills.

The subjects for the study were 56 members of a seventh grade class in a rural central school. These students had taken the 1960 S-B and the Iowa Test of Basic Skills (Form 2) in third grade, and the Lorge-Thorndike Intelligence Tests and Iowa Tests of Basic Skills (Form 1) in grade 6.

The authors concluded that:

1. The 1960 S-B is a valid measure of "scholastic ability;
2. Group mental tests and comprehensive achievement tests lead to the same prediction;
3. For this small sample, at least, the two IQ tests used could be considered interchangeable.

Deloria, Dennis. "The Stanford-Binet in the Ypsilanti Curriculum. Why?" High Scope Educational Research Foundation.

The author evaluates the Stanford-Binet Intelligence Scale as he defends its use as the key instrument used to assess the affects of the Ypsilanti Cognitively-Oriented Follow Through Curriculum.

The following points are made:

1. The Stanford-Binet (S-B) exhibits high predictive validity for later success in school;

2. The S-B excells in reliability;

3. The S-B measures present ability, not inborn capacity;

4. Though the S-B was normed using middle class white persons, it does predict success in American schools as they currently exist better than any comparable test.

5. The cultural bias of the S-B can be diminished by considering change scores rather than absolute scores;

6. An intelligence test is really only a special kind of achievement test;

7. The variability of S-B scores remains essentially constant as children progress in age.

Eagle, N. and Harris, A.S. "Interaction of Race and Test on Reading Performance Scores." Journal of Educational Measurement, 1966, 6(3), 131-135.

This study examines the relationship between race and performance on two nationally standardized reading tests. The appropriate reading tests of the Iowa Tests of Basic Skills and Metropolitan Achievement Battery were administered to all fourth and sixth grade students in all elementary schools of an urban school district near New York City. Although white pupils earned higher scores than non-white pupils on both tests, the Metropolitan produced significantly greater differences between the races than the Iowa, at both grade levels. White upper elementary children were "favored" by the Metropolitan and Negro children were "favored" by the Iowa.

The authors feel that this finding is of importance since, if an interaction exists between social class or ethnic group and reading test, knowledgeable administrators could use combinations of tests to demonstrate either (a) maximum school reading performance for public relations purposes, or (b) minimal school reading performance, possibly to qualify for more Federal or state compensatory program aid.

Feldman, S. and Weiner, M. "The Use of a Standardized Reading Achievement Test with Two Levels of Socio-Economic Status Pupils." The Journal of Experimental Education, 1964, 32(3), 269-274.

The authors are concerned that, while most tests are standardized for normally distributed populations, the populations most often actually tested are homogeneous. Lower and middle socio-economic level fifth grade students were tested with the Gates Advanced Primary Reading Test and the Gates Basic Reading Test, Level of Comprehension subtest.

The purpose of this study was an analysis of these two tests in an effort to determine comparability and the effect of socio-economic level on test scores. After meeting several difficulties in attempting to use such heterogeneously standardized tests for evaluating the homogeneous groups of subjects, the authors made the following recommendations concerning the use of standardized tests in assessing the reading skills of students in different socio-economic levels:

1. Even though test publishers state that grade scores are comparable in adjacent tests in a battery of reading achievement tests, two points must be considered: (a) the highest and lowest possible grade scores for each test; and (b) whether both tests may be measuring different reading skills. When two different ability groups are given the test, then factors a and b above become even more crucial.

2. Tests should be tried out with a small pilot-group before any large-scale test administration is undertaken.

3. If the purpose of the test user is to distinguish among the pupils' skills measured by the test, then the use of a standardized test with a homogeneous group will very likely not provide the desired results.

Himelstein, P. "Research With the Stanford-Binet, Form L-M: The First Five Years." Psychological Bulletin, 1966, 65(3), 156-164.

This article reviews the first five years of research with the 1960 revision of the Stanford-Binet Intelligence Scale (1960-1965). The major areas considered by the reviewer are validity, reliability, sensitivity to extra-test influences, performance of southern Negro children, influence of socio-economic status, diagnostic value, and brief forms.

A major problem in the studies reviewed was the population used; frequently the sample was drawn from retarded or gifted children. This fact decreased the generalizability of the results.

Almost all of the validity studies were concerned with concurrent validity in which the Stanford-Binet was correlated with another test, often a test that just happened to be in the patient's file. The author reports correlations between the Stanford-Binet and the Wechsler Intelligence Scale for Children, the Wechsler Adult Intelligence Scale, the Draw-A-Man Test, the Peabody Picture Vocabulary Test, the Kuhlman Test of Mental Development, the Vineland Social Maturity Scale, the Gesell Developmental Schedule, the California Test of Mental Maturity, the Sequential Test of Educational Progress, the California Achievement Test, and teacher ratings.

The studies revealed little concern for the reliability of the test, the effects of various incentives, and the influence of situational variables, such as "warm" versus "cold" examiners.

In general, the use of the Stanford-Binet short form was found to slightly underestimate the full-scale score. The Stanford-Binet has not been widely employed as a diagnostic tool.

Hirshoren, A. "A Comparison of the Predictive Validity of the Revised Stanford-Binet Intelligence Scale and the Illinois Test of Psycholinguistic Abilities," Exceptional Children, 1969, 35(3), 517-521.

The Illinois Test of Psycholinguistic Ability (ITPA) was developed as a clinical diagnostic test to be used with young children for the identification of specific language disabilities.

The problem of the present study was to determine the predictive validity of the ITPA Total Language Score and of the subtest scores, as well as the predictive validity of the Revised Stanford-Binet IQ, all obtained at the beginning of kindergarten, for school achievement at the beginning of the second grade. For the study, 40 Caucasion children from a heterogeneous neighborhood were tested with the ITPA, the 1960 revision of the Stanford-Binet, and the California Achievement Test (CAT), Lower Primary Form.

The results indicate that the Total Language Score of the ITPA, at least at the kindergarten level, is as valid a predictor as is the Stanford-Binet IQ for school achievement two years later. Performance on the ITPA subtests at the automatic-sequential level is significantly related to performance on all seven of the achievement variables from the CAT. At the representational level of the ITPA, the various subtests are significantly related to various of the CAT subtests. Using stepwise multiple correlation with the ITPA subtests increases the predictive validity of the ITPA. Including the Stanford-Binet IQ as a predictor variable does not improve upon the predictions of future achievement obtained by using the ITPA subtests alone. The ITPA has the added advantage of noting specific areas of language deficits so that remedial programs for individual children may be developed.

Johnson, G.B. "Factors to be Considered in the Interpretation of Intelligence-Test Scores." The Elementary School Journal, 1953, 1-8.

"There are three sources of potential error to be considered in the interpretation of an intelligence quotient. These are to be found in external factors, which relate to the test itself; in internal factors, which include organic and emotional aspects of the individual being tested; and in factors which may be both external and internal and which include the effect of the individual's environment in general." From the detailed analysis of these sources of error, the author concludes:

"1. Select the tests which most nearly suit your purpose in testing. To make the selection requires a knowledge of individual, group, verbal and performance tests and their statistical foundations.

2. Select the tests which have been demonstrated to be both valid and reliable in testing situations similar to your own. This requires a knowledge of the individual and group differences in your situation. Even then, do not put too much credence in the results of one test.

3. If you use standard norms, realize their limitations. If you find them inadequate, prepare local norms.

4. Whichever test you choose, follow closely the directions for administering it. Practice on yourself often before actually administering it.

5. If the test is not machine scored, score it yourself twice.

6. Administer the test under optimum physical and psychological conditions.

7. With both individual and group tests, observe the pupils for factors which may be significant to efficiency of performance, hence to score interpretation.

8. Condition the testees for, not against, intelligence tests. An initial approach in this direction should be to provide pre-service and in-service training for teachers.

9. Understand that intelligence is a composite of factors, in some of which one person may be relatively strong and others relatively weak.

10. Realize that, even with all these precautions, the intelligence test is still in an early phase of development and that the intelligence quotient cannot be considered as precise, as consistent or as true a value as in a unit of measure from the thermometer or the weighing scale."

Knief, L.M. and Stroud, J.B. "Intercorrelations Among Various Intelligence, Achievement and Social Class Scores." Journal of Educational Psychology, 1959, 50(3), 117-120.

In order to obtain information on the relationship of social class to intelligence and achievement test scores, a group of fourth grade students in a Midwestern community were tested over a period of several months. The tests administered were (1) Lorge-Thorndike Intelligence Tests, Verbal; (2) Lorge-Thorndike Intelligence Tests, Nonverbal (Level 3 in both cases); (3) Davis-Eells Games (Elementary Level); (4) Raven's Progressive Matrices; (5) Iowa Test of Basic Skills; and (6) the Warner Index of Status Characteristics.

The following intercorrelations were reported:

| Variable | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | | .309 | .304 | .323 | .179 | .340 |
| 2 | | | .571 | .645 | .405 | .568 |
| 3 | | | | .709 | .437 | .839 |
| 4 | | | | | .521 | .683 |
| 5 | | | | | | .450 |
| 6 | | | | | | |

The authors concluded that their study gave little justification for the use of the Davis-Eells games, Lorge-Thorndike Non-Verbal test, and the Progressive Matrices in conjunction with the Lorge-Thorndike Verbal test for general prediction purposes. However, this information does not negate their usefulness in individual diagnosis.

Powell, W.R. "Gates-MacGinitie Reading Tests." Journal of Educational Measurement, 1969, 6(2), 114-116.

As compared with other general reading tests, the Gates-MacGinitie Reading Tests would provide usable data on achievement in comprehension, vocabulary, and speed. It would be of limited value if information about specific reading sub-skills were needed.

The earlier Gates tests were criticized due to use of one instrument to measure a wide range of grades (3-10). It was challenged as being inappropriate for the highest and lowest levels tested. By breaking down the test into different levels this situation has been remedied. Whether the breakdown has gone too far to be of practical value remains to be seen. Some question also arises as to the speededness of the vocabulary and comprehension tests. Although these tests are considered power tests, 15-25% of the pupils were unable to complete them in the allotted time.

The test manual does not indicate the proper interpretation of the obtained score for translation into classroom practice. The implication is given that the grade score obtained by a student is the instructional reading level for him. This implication is not supported by other evidence and the absence of caution and clarification on the matter to the teacher lessens the effectiveness of use of the test results. While not peculiar to just the Gates-MacGinitie tests, survey or general reading tests tend to produce frustration or prefrustration level scores. An explanation of this interpretation relating the obtained test scores to functional reading levels (independent, instructional, frustration) needs a full discussion in the manual. Then, and only then, will teacher estimates and test scores become more compatible. Even with limitations, the Gates-MacGinitie Reading Tests will undoubtedly continue to enjoy widespread use which was established by the earlier editions.

Rainey, R.G. "A Study of Four School-Ability Tests." The Journal of Experimental Education, 1965, 33(4), 305-319.

The purpose of this study was the evaluation of four school-ability tests: the 1963 versions of the Kuhlman-Anderson, the California Short-Form Test of Mental Maturity, the SRA Primary Mental Abilities Test, and the 1957 form of the Lorge-Thorndike Intelligence Tests. In addition, the Iowa tests of Basic Skills (I.T.B.S.) were administered. The subjects were 2nd, 4th and 7th grade pupils and were designated as belonging to either high, middle or low socio-economic status.

The authors then summarized their results in a series of statements about each test.

The Lorge-Thorndike Test was found to be easy to administer at all levels, from both the students' and testers' viewpoint. The stability of the mean and standard deviation was good at all grade levels in the middle SES area, but poor in the high and low areas. The verbal sections of the test correlated highly with the reading, vocabulary and other language related Iowa Tests; correlations with arithmetic were good, but other tests had better results. The non-verbal section of the test evidenced poor correlations with the Iowa tests when compared with the results obtained from the verbal and total test. The means were highest in the highest SES area and lowest in the low SES area.

The Kuhlman-Anderson Test was easy to administer at all levels and had the added advantage of being short enough to administer during one class period at all grade levels. The mean and standard deviation stability was excellent at all levels and in all SES areas. This test produced the best correlations with the I.T.B.S. in each SES area. The similarity between the means in the high and middle SES areas constituted the major inconsistency of the test.

The California Test of Mental Maturity was not difficult to administer, however, the mean and standard deviation was good only in the high SES area. The language section of the test correlated highly with the verbal subtests of the I.T.B.S. Correlations with arithmetic were higher on the language section than the non-language portions.

The Primary Mental Abilities Test was easy to administer; however, it was the longest of the four tests at each level. The mean and standard deviation stability was the most inconsistent on this test and the greatest variability in all SES areas was produced by this test. The verbal sections of this test correlated well with the verbal subtests of the I.T.B.S. and the total score correlated consistently with all sections of the Iowa Tests. However, none of the Iowa subtests were correlated with the space relations and perceptual speed tests.

As a general conclusion, the authors stated that no one of the tests considered was proven exceptionally inferior or superior.

Stake, R.E. and Hastings, J.T. "The Stanford Achievement Test Battery," Personnel and Guidance Journal, 1964, 43(10), 178-184.

The 1964 edition of the Stanford Achievement Test Battery is stratified in five levels for grades 1-9. At each level, verbal comprehension and skills in rhetoric and quantification are emphasized. For younger pupils there is considerable attention to phonics; for older pupils there are many items devoted to concepts taught in general science, social studies, and advanced arithmetic classes. Determination of whether or not the tests contain appropriate content is in large part a responsibility of local school officials. The Manuals for the Stanford Achievement Tests are well organized to help local officials decide on the appropriateness of content.

The 1964 batteries are longer than the 1963 batteries and the items for the 1964 Edition seem to be more difficult. There is some evidence of the influence of the new curricula, especially in science and mathematics, but school people who have supported innovation in curriculum are likely to have reservations about the item content and the emphasis on grade-equivalent scores.

For the teacher who has to score his own tests, scoring aids packaged with the Stanford Tests are very helpful. The reliability of the sub-tests appears to be on a par with previous edition and they are generally as high as those for any of the currently available, standardized achievement tests. For "knowledge of validity" the authors include a description of the sampling of content, and they put all emphasis upon the user's judgment regarding appropriateness of content for his school.

The arrangement for profiling students' scores on the front of the test booklet or on a separate sheet is attractive and the method of profiling is easy to execute. However, the user should be wary since the use of stanines and the size of intercorrelations may aid over-interpretation.

It seems safe to conclude that if local instructors endorse the content coverage, the Stanford Tests will do as effective a job of measuring elementary school achievement as any standardized battery currently available.